AA

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 August 2005 (04.08.2005)

PCT

(10) International Publication Number
WO 2005/070961 A1

(51) International Patent Classification[7]: C07K 14/705, G01N 33/68

(21) International Application Number: PCT/US2005/002333

(22) International Filing Date: 24 January 2005 (24.01.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/763,644 22 January 2004 (22.01.2004) US

(71) Applicant *(for all designated States except US)*: **XENOPORT, INC.** [US/US]; 3410 Central Expressway, Santa Clara, California 95051-0703 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **ZERANGUE, Noa** [US/US]; 3330 La Mesa Drive #12, San Carlos, California 94070 (US). **PADDON, Christopher, J.** [US/US]; 1046 Banyan Way, Pacifica, California 94044 (US).

(74) Agents: **LIEBESCHUETZ, Joe** et al.; TOWNSEND AND TOWNSEND AND CREW LLP, Two Embarcadero Center, 8th Floor, San Francisco, California 94111-3834 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HUMAN ORGANIC SOLUTE TRANSPORTERS

(57) Abstract: The invention provides several new transporter proteins, nucleic acids encoding them and methods of using the transporters to screen agents, conjugates or conjugate moieties, linked or linkable to agents, for capacity to be transported as substrates through the transporters. The invention also provides methods of treatment involving oral delivery of agents that either alone, or as a result of linkage to a conjugate moiety, are substrates of one of the transporter.

WO 2005/070961 A1

Human Organic Solute Transporters

CROSS-REFERENCE TO RELATED APPLICATION

This application is a continuation of USSN 10/763,644 filed January 22, 2004.

BACKGROUND OF THE INVENTION

[0001] Recent advances in the pharmaceutical industry have resulted in the formation of an increasing number of potential therapeutic agents. However, formulating the compounds for effective oral bioavailability has proven difficult because of problems associated with uptake and high susceptibility to metabolic enzymes.. Natural transporter proteins are involved in the uptake of various molecules into and/or through cells. One strategy for the delivery is to identify pharmacological agents that are, or can be modified to be, substrates for transport proteins.

[0002] In general, two major transport systems exist: solute carrier-mediated systems and receptor mediated systems. Carrier-mediated systems use transport proteins that are anchored to the cell membrane, typically by a plurality of membrane-spanning loops and function by transporting their substrates via an energy-dependent flip-flop or other mechanism, exchange and other facilitative or equilibrative mechanisms. Carrier-mediated transport systems are involved in the active or non-active, facilitated transport of many important nutrients such as vitamins, sugars, and amino acids, as well as xenobiotic compounds. The carrier systems result in result in transport into the enterocytes from blood or lumen, and across the epithelial cell layer from lumen into blood (absorption) or blood to lumen (secretion). Carrier-mediated transporters are also present in organs such as liver and kidney, in which the proteins are involved in the excretion or re-absorption of circulating compounds.

[0003] Receptor-mediated transport systems differ from the carrier-mediated systems in that receptors usually span the cell membrane only a single time. Furthermore, substrate binding triggers an invagination and encapsulation process that results in the formation of various transport vesicles to carry the substrate (and sometimes other molecules) into and through the cell. This process of membrane deformations that result in the internalization of certain substrates and their subsequent targeting to certain locations in the cytoplasm is generally referred to as endocytosis.

[0004] Polar or hydrophilic compounds are typically poorly absorbed through an animal's intestine as there is a substantial energetic penalty for passage of such

compounds across the lipid bilayers that constitute cellular membranes. Many nutrients that result from the digestion of ingested foodstuffs in animals, such as amino acids, di- and tripeptides, monosaccharides, nucleosides and water-soluble vitamins, are polar compounds whose uptake is essential to the viability of the animal. For these substances there exist specific mechanisms for active transport of the solute molecules across the apical membrane of the intestinal epithelia. This transport is frequently energized by co-transport of ions down a concentration gradient.

[0005] An organic solvent transporter and an ancillary membrane protein, termed Ostα and β, have recently been cloned from skate, Wang *et al.*, *PNAS* 98, 9431-9436 (2001). Ostα encodes a protein of 352 amino acids and seven putative transmembrane domains. Ostβ encodes a protein of 182 amino acids with at least one and perhaps two transmembrane domains. Xenopus oocytes transfected with nucleic acids encoding Ostα and β were reported to transport labeled taurocholate. Wang *et al.*, report searching databases for sequences showing significant sequence identity with Ostα and β but did not find any such sequences.

[0006] PCT/US03/32087, filed on October 8, 2003 and USSN 60/417,298 filed October 8, 2002 directed to related subject matter are incorporated by reference in their entirety for all purposes.

SUMMARY OF THE INVENTION

[0007] The present invention provides new transporter polypeptides in isolated form. Isolated polypeptides of the invention are at least 80% identical to an amino acid sequence as set forth in a sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, and 8, over a region of at least 40 amino acids in length when compared using the BLASTP algorithm with a wordlength (W) of 3, and the BLOSUM62 scoring matrix. Some polypeptides of the invention specifically binds to an antibody that specifically binds to a polypeptide selected from the group consisting of SEQ ID NOS: 2, 4, 6, and 8. SEQ ID NOS: 2, 4, 6 and 8 are exemplary amino acid sequences of transporter proteins.

[0008] The invention also provides isolated nucleic acids having a sequence at least 80% identical to a polynucleotide having a sequence selected from the group consisting of SEQ ID NOS: 1, 3, 5, and 7 over a region at least 100 nucleotides in length when compared using the BLASTN algorithm with a wordlength (W) of 11, M=5, and N=-4. Some isolated nucleic acids of the invention hybridize to a sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, and 7 under high stringency conditions, including 50%

formamide, 5X SSC, 5X Denhardt's solution, 10 mM sodium phosphate, pH 6.5, 100 mg/ml salmon sperm DNA at 42°. SEQ ID NOS: 1, 3, 5, and 7 are exemplary nucleic acids of the invention. Optionally, the isolated nucleic acids are provided as components of a vector.

[0009] The invention further provides screening methods to determine whether an agent, conjugate or conjugate moiety is a substrate of a transporter. Some methods include providing a cell expressing a nucleic acid as described above in the outer membrane of the cell, contacting the cell with an agent, conjugate moiety or conjugate, and determining whether the agent, conjugate moiety or conjugate passes through the transporter. In some methods, the transporter encoded by a nucleic acid has the sequence of SEQ ID NO:2. In some methods, the cells used for expression of the transporter are Chinese hamster ovary cells, human embryonic kidney cells or oocytes.

[0010] The invention further provides methods to determine whether an agent, conjugate moiety or conjugate binds a transporter. The transporter, which has a sequence with at least 80% sequence identity to an amino acid sequence as set forth in a sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, and 8, over a region of at least 40 amino acids in length when compared using the BLASTP algorithm with a wordlength (W) of 3, and the BLOSUM62 scoring matrix, is contacted with an agent, conjugate moiety or conjugate and the presence or absence of binding between the transporter and the agent, conjugate moiety or conjugate is detected.

[0011] The invention further provides a conjugate comprising an agent linked to a conjugate moiety for a transporter polypeptide as described above. The conjugate shows a Vmax of at least 1% of taurocholate for the transporter wherein the agent has a pharmaceutical activity without the conjugate moiety, and the conjugate has a greater Vmax for the transporter than the agent without the conjugate moiety.

[0012] The invention further provides methods of manufacturing a pharmaceutical composition, comprising linking an agent to a conjugate moiety to form a conjugate wherein the conjugate is transported by a transporter having an amino acid sequence with at least 80% identity to an amino acid sequence as set forth in a sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, and 8, over a region of at least 40 amino acids in length when compared using the BLASTP algorithm with a wordlength (W) of 3, and the BLOSUM62 scoring matrix, with a Vmax higher than the agent alone, formulating the conjugate with a carrier as a pharmaceutical composition.

[0013] The invention further provides methods of treatment comprising administering to patient a conjugate comprising an agent linked to a conjugate moiety

wherein the conjugate is transported by a transporter having an amino acid sequence with at least 80% identity to an amino acid sequence as set forth in a sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, and 8, over a region of at least 40 amino acids in length when compared using the BLASTP algorithm with a wordlength (W) of 3, and the BLOSUM62 scoring matrix with a Vmax higher than the agent alone. In some methods the conjugate is administered orally to the patient. In other methods, the conjugate is administered intravenously.

[0014] The invention provides methods of screening whether a conjugate or conjugate moiety is a substrate of a transporter. A cell expressing a nucleic acid to produce a transporter in an plasma membrane of the cell, the transporter having an amino acid sequence at least 90 % identical to SEQ ID NO: 2 over the entire length of SEQ ID NO:2, wherein the transporter is coexpressed with a co-transporter having at least 90% sequence identity to SEQ ID NO:20 over the entire length of the SEQ ID NO:20 is provided. The cell is contacted with a conjugate moiety or conjugate. Whether the conjugate moiety or conjugate passes through the plasma membrane via the transporter is determined.

[0015] The invention further provides methods of screening agents, conjugate moieties or conjugates as substrate for intestinal transport. A agent, conjugate moiety or conjugate is screened for capacity to be a substrate for a known transporter that effects transport through an apical plasma membrane of epithelial cells lining the intestine. The agent, conjugate moiety or conjugate is also screened for capacity to be a substrate for a transporter having the amino acid sequence of SEQ ID NO:2.

[0016] The invention provides methods of screening agents, conjugates or conjugate moieties for capacity to agonize or antagonize a transporter. A cell expressing a transporter having an amino acid sequence at least 90 % identical to SEQ ID NO: 2 over the entire length of SEQ ID NO:2, wherein the transporter is coexpressed with a co-transporter having at least 90% sequence identity to SEQ ID NO:20 over the entire length of the SEQ ID NO:20, the transporter being situated in the plasma membrane of the cell; is contacted with an agent, conjugate or conjugate moiety and a known substrate of the transporter. Whether the agent agonizes or antagonizes uptake of the known substrate into the cell in comparison with a control cell expressing the transporter contacted with known substrate without the agent, conjugate or conjugate moiety is determined.

[0017] The invention further provides a conjugate comprising an agent linked to a conjugate moiety for a transporter having the amino acid sequence SEQ ID NO: 2, wherein the transporter is coexpressed with a co-transporter having the amino acid sequence

SEQ ID NO:20, wherein the conjugate shows a Vmax of at least 1% of taurocholate for the transporter, wherein the agent has a pharmaceutical activity without the conjugate moiety, and the conjugate has a greater Vmax for the transporter than the agent without the conjugate moiety.

[0018] The invention further provides a method of manufacturing a pharmaceutical composition, comprising; linking an agent to a conjugate moiety to form a conjugate wherein the conjugate is transported by a transporter having the amino acid sequence SEQ ID NO: 2 over the entire length of SEQ ID NO:2, wherein the transporter is coexpressed with a co-transporter having the amino acid sequence SEQ ID NO:20, with a higher Vmax than the agent alone; and formulating the conjugate with a carrier as a pharmaceutical composition.

[0019] The invention further provides the use of a conjugate comprising an agent linked to a conjugate moiety wherein the conjugate is transported by a transporter having the amino acid sequence SEQ ID NO: 2 over the entire length of SEQ ID NO:2, wherein the transporter is coexpressed with a co-transporter having the amino acid sequence SEQ ID NO:20, with a higher Vmax than the agent alone in the manufacture of a medicament.

[0020] The invention further provides for the use of a drug covalently linked to a conjugate moiety through linkage cleavable in vivo to form a conjugate in the manufacture of a medicament for oral administration, wherein the conjugate is transported by a transporter having an amino acid sequence of SEQ ID NO: 2, wherein the transporter is coexpressed with a co-transporter having an amino acid sequence of SEQ ID NO:20, the transporter being situated in the plasma membrane of epithelial cells lining the intestinal lumen, wherein the conjugate has a Vmax of at least 1% of taurocholate for the transporter, wherein oral bioability of the agent due to uptake of the conjugate is improved relative to the oral bioavailabilityof the agent administered without the conjugate moiety.

DEFINITIONS

[0021] The term "nucleic acid" refers to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, encompasses known analogues of natural nucleotides that hybridize to nucleic acids in a manner similar to naturally-occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence includes the complementary sequence thereof. A

"subsequence" refers to a sequence of nucleotides or amino acids that comprise a part of a longer sequence of nucleotides or amino acids (*e.g.*, a polypeptide), respectively.

[0022] A "probe" is a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation, thus forming a duplex structure. The probe binds or hybridizes to a "probe binding site." A probe may include natural (*i.e.*, A, G, C, or T) or modified bases (7-deazaguanosine, inosine, *etc.*). A probe can be an oligonucleotide which is a single-stranded DNA. Oligonucleotide probes can be synthesized or produced from naturally occurring polynucleotides. In addition, the bases in a probe can be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages (see, for example, Nielsen *et al.*, *Science* 254, 1497-1500 (1991)). Some probes may have leading and/or trailing sequences of noncomplementary flanking a region of complementarity.

[0023] The terms "polypeptide," "peptide" and "protein" are used interchangeably to refer to a polymer of amino acid residues. The term also applies to amino acid polymers in which one or more amino acids are chemical analogues of corresponding naturally-occurring amino acids.

[0024] The term "operably linked" refers to functional linkage between a nucleic acid expression control sequence (such as a promoter, signal sequence, or array of transcription factor binding sites) and a second polynucleotide, wherein the expression control sequence affects transcription and/or translation of the second polynucleotide.

[0025] A "heterologous sequence" or a "heterologous nucleic acid," as used herein, is one that originates from a source foreign to the particular host cell, or, if from the same source, is modified from its original form. Thus, a heterologous gene in a prokaryotic host cell includes a gene that, although being endogenous to the particular host cell, has been modified. Modification of the heterologous sequence can occur, *e.g.*, by treating the DNA with a restriction enzyme to generate a DNA fragment that is capable of being operably linked to the promoter. Techniques such as site-directed mutagenesis are also useful for modifying a heterologous nucleic acid.

[0026] The term "recombinant" when used with reference to a cell indicates that the cell replicates a heterologous nucleic acid, or expresses a peptide or protein encoded by a heterologous nucleic acid. Recombinant cells can contain genes that are not found within the native (non-recombinant) form of the cell. Recombinant cells can also contain

genes found in the native form of the cell wherein the genes are modified and re-introduced into the cell by artificial means. The term also encompasses cells that contain a nucleic acid endogenous to the cell that has been modified without removing the nucleic acid from the cell; such modifications include those obtained by gene replacement, site-specific mutation, and related techniques.

[0027] The term "isolated," "purified" or "substantially pure" means an object species that has been enriched or separated from the components in its native environment. Thus, a nucleic acid that is being recombinantly expressed in vitro is isolated notwithstanding that the nucleic acid is surrounded by other cellular components. The term may also indicate the an object species is the predominant macromolecular species present (*i.e.*, on a molar basis it is more abundant than any other individual species in the composition), and preferably the object species comprises at least about 50 percent (on a molar basis) of all macromolecular species present. Generally, an isolated, purified or substantially pure composition will comprise more than 80 to 90 percent of all macromolecular species present in a composition. Most preferably, the object species is purified to essential homogeneity (*i.e.*, contaminant species cannot be detected in the composition by conventional detection methods) wherein the composition consists essentially of a single macromolecular species.

[0028] The term "complementary" means that one nucleic acid is identical to, or hybridizes selectively to, another nucleic acid molecule. Selectivity of hybridization exists when hybridization occurs that is more selective than total lack of specificity. Typically, selective hybridization will occur when there is at least about 55% identity over a stretch of at least 14-25 nucleotides, preferably at least 65%, more preferably at least 75%, and most preferably at least 90%. Preferably, one nucleic acid hybridizes specifically to the other nucleic acid. See M. Kanehisa, *Nucleic Acids Res. 12:203 (1984).*

[0029] The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptides, refer to two or more sequences or subsequences that are the same or have a specified percentage of nucleotides or amino acid residues that are the same, when compared and aligned for maximum correspondence, as measured using a sequence comparison algorithm such as those described below for example, or by visual inspection.

[0030] The phrase "substantially identical," in the context of two nucleic acids or polypeptides, refers to two or more sequences or subsequences that have at least 75%, preferably at least 85%, more preferably at least 90%, 95% or higher nucleotide or amino acid residue identity, when compared and aligned for maximum correspondence, as measured

using a sequence comparison algorithm such as those described below for example, or by visual inspection. Preferably, the substantial identity exists over a region of the sequences that is at least about 40-50 residues in length, preferably over a longer region than 50 amino acids, more preferably at least about 90-100 residues, and most preferably the sequences are substantially identical over the full length of the sequences being compared, such as the coding region of a nucleotide for example.

[0031] For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

[0032] Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by visual inspection (see generally Ausubel *et al.*, *supra*).

[0033] Another example of algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra.*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to

calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. For identifying whether a nucleic acid or polypeptide is within the scope of the invention, the default parameters of the BLAST programs are suitable. The BLASTN program (for nucleotide sequences) uses as defaults a word length (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a word length (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix. The TBLATN program (using protein sequence for nucleotide sequence) uses as defaults a word length (W) of 3, an expectation (E) of 10, and a BLOSUM 62 scoring matrix. (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)).

[0034] In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, *e.g.*, Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

[0035] Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions. "Bind(s) substantially" refers to complementary hybridization between a probe nucleic acid and a target nucleic acid and embraces minor mismatches that can be accommodated by reducing the stringency of the hybridization media to achieve the desired detection of the target polynucleotide sequence. The phrase "hybridizing specifically to", refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (*e.g.*, total cellular) DNA or RNA.

[0036] The term "stringent conditions" refers to conditions under which a probe will hybridize to its target subsequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are

selected to be about 5° C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. (As the target sequences are generally present in excess, at Tm, 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60° C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide.

[0037] A further indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid, as described below. The phrases "specifically binds to a protein" or "specifically immunoreactive with," when referring to an antibody refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, a specified antibody binds preferentially to a particular protein and does not bind in a significant amount to other proteins present in the sample. Specific binding to a protein under such conditions requires an antibody that is selected for its specificity for a particular protein. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See, *e.g.*, Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity.

[0038] "Conservatively modified variations" of a particular polynucleotide sequence refers to those polynucleotides that encode identical or essentially identical amino acid sequences, or where the polynucleotide does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given polypeptide. For instance, the codons CGU, CGC, CGA, CGG, AGA, and AGG all encode the amino acid arginine. Thus, at every position where an arginine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such

nucleic acid variations are "silent variations," which are one species of "conservatively modified variations." Every polynucleotide sequence described herein which encodes a polypeptide also describes every possible silent variation, except where otherwise noted. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine) can be modified to yield a functionally identical molecule by standard techniques. Accordingly, each "silent variation" of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

[0039] A polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. A "conservative substitution," when describing a protein, refers to a change in the amino acid composition of the protein that does not substantially alter the protein's activity. Thus, "conservatively modified variations" of a particular amino acid sequence refers to amino acid substitutions of those amino acids that are not critical for protein activity or substitution of amino acids with other amino acids having similar properties (*e.g.*, acidic, basic, positively or negatively charged, polar or non-polar, *etc.*) such that the substitutions of even critical amino acids do not substantially alter activity. Conservative substitution tables providing functionally similar amino acids are well-known in the art. See, *e.g.*, Creighton (1984) Proteins, W.H. Freeman and Company. In addition, individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids in an encoded sequence are also "conservatively modified variations."

[0040] Allelic variants of a gene refer to variant forms of the same gene between different individuals of the same species. Cognate forms of a gene refers to variation between structurally and functionally related genes between species. For example, the human gene showing the greatest sequence identity and functional related to a mouse gene is the human cognate form of the mouse gene.

[0041] The term "naturally-occurring" as applied to an object refers to the fact that an object can be found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism that can be isolated from a source in nature and which has not been intentionally modified by humans in the laboratory is naturally-occurring.

[0042] The term "antibody" refers to a protein consisting of one or more polypeptides substantially encoded by immunoglobulin genes or fragments of immunoglobulin genes. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon and mu constant region genes, as well as myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda.

Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

[0043] The term "patient" includes human and veterinary subjects.

[0044] The phrases "specifically binds" when referring to a protein or "specifically immunoreactive with" when referring to an antibody, refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biologics. Thus, under designated conditions, a specified ligand binds preferentially to a particular protein and does not bind to a significant extentto other proteins present in the sample. A molecule such as antibody that specifically binds to a protein often has an association constant of at least 10 M^{-1}, $10^6 M^{-1}$ or $10^7 M^{-1}$, preferably $10^8 M^{-1}$ to $10^9 M^{-1}$, and more preferably, about $10^{10} M^{-1}$ to $10^{11} M^{-1}$ or higher. However, some substrates of a transporter have much lower affinities of the order of $10\text{-}10^3$ M^{-1} and yet the binding can still be shown to be specific. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See, *e.g.*, Harlow and Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity.

[0045] A "transport protein" is a protein that has a direct or indirect role in transporting a molecule into and/or through a cell. The term includes, for example, membrane-bound proteins that recognize a substrate and effects its entry into, or exit from a cell by a carrier-mediated transporter or by receptor-mediated transport. These proteins are sometimes referred to as transporter proteins. The term also includes intracellularly expressed proteins that participate in trafficking of substrates through or out of a cell. The term also includes proteins or glycoproteins exposed on the surface of a cell that do not directly transport a substrate but bind to the substrate holding it in proximity to a receptor or transporter protein that effects entry of the substrate into or through the cell. Examples of carrier proteins include: the intestinal and liver bile acid transporters, dipeptide transporters, oligopeptide transporters, simple sugar transporters (*e.g.*, SGLT1), phosphate transporters, monocarboxcylic acid transporters, P-glycoprotein transporters, organic anion transporters (OATP), and organic cation transporters. Examples of receptor-mediated transport proteins include: viral receptors, immunoglobulin receptors, bacterial toxin receptors, plant lectin

receptors, bacterial adhesion receptors, vitamin transporters and cytokine growth factor receptors.

[0046] A "substrate" of a transport protein is a compound whose uptake into or passage through a cell is facilitated by the transport protein.

[0047] The term "ligand" of a transport protein includes substrates and other compounds that bind to the transport protein without being taken up or transported through a cell. Some ligands by binding to the transport protein inhibit or antagonize uptake of the substrate or passage of substrate through a cell by the transport protein. Some ligands by binding to the transport protein promote or agonize uptake or passage of the compound by the transport protein or another transport protein. For example, binding of a ligand to one transport protein can promote uptake of a substrate by a second transport protein in proximity with the first transport protein.

[0048] The term "agent" is used to describe a compound that has or may have a pharmacological activity. Agents include compounds that are known drugs, compounds for which pharmacological activity has been identified but which are undergoing further therapeutic evaluation, and compounds that are members of collections and libraries that are to be screened for a pharmacological activity.

[0049] An agent is "orally active" if it can exert a pharmaceutical activity when administered via an oral route.

[0050] A "conjugate moiety" refers to a compound or part of a compound that does not itself have pharmacological activity but which can be linked to an agent to form a conjugate that does have pharmacological activity. Typically, the agent has pharmacologic activity without the conjugate moiety. The conjugate moiety facilitates therapeutic use of the agent by promoting uptake of the agent via a transporter. A conjugate moiety can itself be a substrate for a transporter or can become a substrate when linked to a compound (*e.g.*, valacyclovir). Thus, a conjugate moiety formed from a compound and a conjugate moiety can have higher uptake activity than either the compound or moiety alone.

[0051] A "pharmacological" activity means that an agent exhibits an activity in a screening system that indicates that the agent is or may be useful in the prophylaxis or treatment of a disease. The screening system can be in vitro, cellular, animal or human. Agents can be described as having pharmacological activity notwithstanding that further testing may be required to establish actual prophylactic or therapeutic utility in treatment of a disease.

[0052] Vmax and Km of a compound for a transporter are defined in accordance with convention. Vmax is the number of molecules of compound transported per second at saturating concentration of the compound. Km is the concentration of the compound at which the compound is transported at half of Vmax. In general, a high value of Vmax is desirable for a substrate of a transporter. A low value of Km is desirable for transport of low concentrations of a compound, and a high value of Km is desirable for transport of high concentrations of a compound. Vmax is affected both by the intrinsic turnover rate of a transporter (molecules/transporter protein) and transporter density in plasma membrane which depends on expression level. For these reasons, the intrinsic capacity of a compound to be transported by a particular transporter is usually expressed as the ratio Vmax of the compound/Vmax of a control compound known to be a substrate for the transporter.

DETAILED DESCRIPTION OF THE INVENTION

I. Introduction

[0053] The invention provides several new transporter protein, nucleic acids encoding them and methods of using the transporters to screen agents, conjugates or conjugate moieties, linked or linkable to agents, for capacity to be transported as substrates through the transporters. The invention also provides methods of treatment involving oral delivery of agents that either alone, or as a result of linkage to a conjugate moiety, are substrates of one of the transporter.

II. The transporter proteins

[0054] The invention provides nucleic acid and amino acid sequences for several human organic solvent transporter proteins, designated hOst-1, hOst-2, hOst3 and hOst-4. One transporter protein, hOst-4, is believed to be the human cognate form of the skate Ostα gene. The nucleic acid and amino acid sequences of hOst-4 are designated SEQ. ID. NOS: 1 and 2. Ost4 shows about 38% sequence identity with the skate Ostα protein sequence (SEQ ID NO:17)) (see Table 1).

Table 1: Pairwise Percent Identity of Peptide Sequences

	hOST1	hOST2	hOST3	hOST4	skate OSTα

hOST1	100%	64%	34%	17%	18%
hOST2	64%	100%	33%	17%	19%
hOST3	34%	33%	100%	19%	18%
hOST4	17%	17%	19%	100%	38%
skateOSTα	18%	19%	18%	38%	100%

[0055] The theoretical location of transmembrane domains is shown in Table 2 below. The numbers refer to the location of amino acid residues, with the initiator methionine as the first amino acid. Transmembrane domains were allocated by alignment of all OST sequences with skate OSTα using ClustalW, and comparison to the transmembrane domains predicted in OST alpha by a Hidden Markov Model (Sonnhammer, *et al.*. A Hidden Markov Model for Predicting Transmembrane Helices in Protein Sequences, *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology*, pages 175-182, Menlo Park, CA, 1998. AAAI Press.). The predicted N-terminus is extracellular, while the C-terminus is intracellular.

[0056] Other structural features of OST genes include potential glycosylation sites, such as at amino acid residues 3-5 of hOST2 and at amino acid residues 203-205 of hOST3. In addition the OST genes contain cysteine residues in the fourth hydrophilic domain. The skate OSTα has a cluster of six cysteine residues between the fourth and fifth transmembrane domains. Likewise, hOST 1 and hOST2 have 2 cysteine residues, while hOST3 has 3 cysteines and hOST4 has 7 cysteine residues.

Table 2:Putative Transmembrane Location

	1	2	3	4	5	6	7
Gene							
hOST1	12-34	47-67	80-101	139-159	173-195	215-237	263-280
hOST2	57-77	90-109	123-144	182-201	216-238	158-280	307-324
hOST3	48-71	83-103	113-134	172-191	206-228	248-270	295-312
hOST4	53-75	88-108	118-139	177-200	214-237	256-278	300-317

[0057] Nucleic acid and amino acid sequences for hOst-1, hOst-2, and hOst3 are designated SEQ. ID. NOS: 3-8 respectively. The percentage sequence identities between the amino acid sequences are shown in Table 1. hOst-1, hOst-2 and hOst-3 are probably nonallelic with each other and hOst-4 but show significant structural and functional relatedness. The relationship is strongest between hOst-1 and hOst-2, which show 64 %

sequence identity with each other. The amino acid sequence of a cotransporter that can be expressed in combination with hOst-1, hOst-2, hOst-3, and particularly, hOst-4 is provided as SEQ ID NO: 19.

[0058] hOst-4 was identified by searching sequence databases for sequences having similarity with skate Ostα. A sequence for hOst-4 was compiled from more than one contig. This sequence was then used to design primers to amplify hOst-4 cDNA. hOst-1, hOst-2 and hOst-3 were also identified by searching databases for sequence similarity with skate Ostα and compiling different parts of the coding regions. The nucleic acid and deduced amino acid sequences provided come directly from the compilation of sequences from the database search without resequencing of cDNAs as was done for hOst-4. Reference to one of the above transporter proteins includes the full-length molecule and other polypeptides having a similar activity. The term also includes mature forms of transporters lacking a signal sequence. The invention also provides variants of exemplified transporter proteins having an amino acid sequence at least 80% identical to an amino acid sequence as set forth in SEQ ID NOS: 2, 4, 6, or 8. More preferably, the variants are at least 85% identical, still more preferably at least 90% or 95% identical to the amino acid sequence of SEQ ID NOS: 2, 4, 6, or 8. The region of similarity between a variant and an exemplary SEQ. ID typically extends over a region of at least 40 amino acids in length, more preferably over a longer region than 40 amino acids such as 50, 60, 70 or 80 amino acids, and most preferably over the full length of the polypeptide. One example of an algorithm that is useful for comparing a polypeptide to a SEQ ID NO. is the BLASTP algorithm; suitable parameters include a word length (W) of 3, and a BLOSUM62 scoring matrix. Variants include allelic variants, splice variants and cognate variants from mammalian species, particularly primates, bovines, canines, felines and rodents.

[0059] Besides substantially full-length polypeptides, the present invention provides for biologically active fragments of the polypeptides. Biological activity may include transport of a substrate that is also transported by the full length protein. Other examples of significant biological activity include antibody binding (*e.g.*, a fragment competes with a full-length sequence as set forth in SEQ ID NOS: 2, 4, 6, or 8) and immunogenicity (*i.e.*, possession of epitopes that stimulate B- or T-cell responses against the fragment. Fragments ordinarily comprise at least 5 contiguous amino acids, typically at least 6 or 7 contiguous amino acids, more typically 8 or 9 contiguous amino acids, usually at least 10, 11 or 12 contiguous amino acids, preferably at least 13 or 14 contiguous amino acids, more preferably at least 16 contiguous amino acids, and most preferably at least 20, 40, 60 or

80 contiguous amino acids. Other examples of subsequences provided by the invention are amino acid sequences wherein 1 to 10 amino acids are removed from the N-terminal of C-terminal end of SEQ ID NOS: 2, 4, 6, or 8.

[0060] Transporter proteins of the invention often share at least one antigenic determinant in common with the amino acid sequence set forth in SEQ ID NOS: 2, 4, 6, or 8. The existence of such a common determinant is evidenced by cross-reactivity of the variant protein with any antibody prepared against the full-length polypeptide. Cross-reactivity may be tested using polyclonal sera against the full-length, but can also be tested using one or more monoclonal antibodies against the full-length protein.

[0061] Some transporter proteins of the invention have a modified polypeptide backbone. Illustrative examples of such modifications include chemical derivatizations of polypeptides, such as acetylations and carboxylations. Modifications also include glycosylation modifications and processing variants of a typical polypeptide. Such processing steps specifically include enzymatic modifications, such as ubiquitinization and phosphorylation. See, *e.g.*, Hershko & Ciechanover, *Ann. Rev. Biochem.* 51:335-364 (1982). Modifications also include substitutions with nonnaturally occurring amino acids, such as D amino acids.

[0062] Expression date for the transporters in a panel of human tissues is provided in Table 3. It can be seen that hOst-1, -2, -3 and -4 are all expressed significantly in the human intestine.

Table 3: mRNA Abundance, transcripts/reaction, normalized to GAPDH

Human Intestine Mucosal Biopsy Samples

	STOMACH	DUODENUM	ILEUM	CECUM	A.COLON	T.COLON	S.COLON	RECTUM	KIDNEY	LIVER	OVARY	TESTIS
OST-1	15218	16789	12214	9644	7730	9050	10163	7957	18188	9743	12498	2367
OST-2	10033	17643	15518	6728	10112	11130	10712	6604	1600	4987	0	750
OST3	1530	6970	9449	8378	8881	9471	10060	1639	1654	2993	N.D.	N.D.
OST4 (OSTalpha)	4866	42267	65347	24009	51725	55623	19137	1242	11058	55863	9937	24683

III. Nucleic Acids

[0063] The present invention further provides isolated and/or recombinant nucleic acids that encode the transporters of the invention or fragments thereof discussed above. The nucleic acids of the invention include naturally occurring, synthetic, and intentionally manipulated polynucleotide sequences (*e.g.*, site directed mutagenesis or use of alternate promoters for RNA transcription). The nucleic acids of the invention also include sequences that are degenerate as a result of the degeneracy of the genetic code.

[0064] The polynucleotide encoding transporters with SEQ ID NOS: 2, 4, 6, or 8 are SEQ ID NOS: 1, 3, 5, and 7 respectively. Also included in the invention are subsequences of the above-described nucleic acid sequences. Such subsequences include, for example, the coding region of SEQ ID NOS: 1, 3, 5, and 7, with or without signal sequences, as well as subsequences that are at least 17, 18, 19, 20, 21, 22, 23, 24 or 25 nucleotides in length.

[0065] The invention also includes polynucleotide sequences that are typically substantially identical to a polynucleotide sequence of SEQ. ID NOS: 1, 3, 5, and 7. For example, the invention includes polynucleotide sequences that are at least about 80% identical to the nucleic acid SEQ ID NO: 1 over a region of at least about 100 nucleotides in length. More preferably, the nucleic acids of the invention are at least 85% identical to the nucleic acid sequence shown in SEQ ID NO: 1, and still more preferably at least 90-95% identical to the nucleic acid sequence of SEQ ID NO: 1 over a region of at least 100 nucleotides.. In some instances, the region of percent identity extends over a longer region such as 125, 150, 175, 200, 225 or 250 nucleotides, or over the full length of the encoding region. To identify nucleic acids of the invention, one can employ a nucleotide sequence comparison algorithm such as are known to those of skill in the art. For example, one can use the BLASTN algorithm. Suitable parameters for use in BLASTN are wordlength (W) of 11, M=5 and N=-4.

[0066] Alternatively, one can identify a nucleic acid of the invention by hybridizing, under stringent conditions, the nucleic acid of interest to a nucleic acid that includes a polynucleotide sequence of SEQ. ID NOS: 1, 3, 5, and 7. The invention also includes nucleic acids which encode a polypeptide which is immunologically cross reactive with SEQ ID NOS: 2, 4, 6, or 8 subsequences thereof.

[0067] Nucleic acid sequences of the present invention can be obtained by methods such as for example, 1) hybridization of genomic or cDNA libraries with probes to detect homologous nucleotide sequences; 2) antibody screening of expression libraries to

detect cloned DNA fragments with shared structural features; 3) various amplification procedures such as polymerase chain reaction (PCR) using primers capable of annealing to the nucleic acid of interest; 4) direct chemical synthesis. Suitable primers for amplification of the coding sequence of the genes from cDNA or DNA are depicted in Table 4 below:

Gene	Primer	SEQ ID NO
hOST1	atggagcagcctgtgttcctgatg	9
	ctagaattcatcatcagagctgagc	10
hOST2	atgagtaatgtctcagggatcctgg	11
	ctacaggtcctccgaggggatcagc	12
hOST3	atgccttgcacttgtacctggagg	13
	tcaggaatccacggatttatctgaag	14
hOST4	atggagccgggcaggacccagataa	15
	ttaggctttgaggttcaagtccagg	16

IV. Production of transporter proteins

[0068] The transporter proteins of SEQ. ID NOS: 2, 4, 6, and 8 and variants thereof can be produced in prokaryotic or eukaryotic host cells by expression of polynucleotides encoding the transporter proteins. DNA sequences are expressed in hosts after the sequences have been operably linked to an expression control sequence in an expression vector. Expression vectors are typically replicable in the host organisms either as episomes or as an integral part of the host chromosomal DNA. Commonly, expression vectors contain selection markers, *e.g.*, tetracycline resistance or hygromycin resistance, to permit detection and/or selection of those cells transformed with the desired DNA sequences (see, *e.g.*, U.S. Patent 4,704,362).

[0069] Typically, the polynucleotide that encodes a polypeptide of the invention is placed under the control of a promoter that is functional in the desired host cell to produce relatively large quantities of a polypeptide of the invention. Ordinarily, the promoter selected depends upon the cell in which the promoter is to be active. Other expression control sequences such as ribosome binding sites, transcription termination sites and the like are also optionally included. Constructs that include one or more of these control sequences are termed "expression cassettes." Accordingly, the invention provides expression cassettes into which the nucleic acids that encode the polypeptides described herein are incorporated for high level expression in a desired host cell. For expression of the polypeptides in

mammalian cells, convenient promoters include CMV promoter (Miller, *et al., BioTechniques 7:980*), SV40 promoter (de la Luma, *et al.*,(1998) *Gene 62:121)*, RSV promoter (Yates, *et al., Nature 313:812*(1985), MMTV promoter (Lee, *et al., Nature 294:228* (1981)).

[0070] Either constitutive or regulated promoters can be used in the present invention. Regulated promoters can be advantageous because the host cells can be grown to high densities before expression of the polypeptides is induced. High level expression of heterologous proteins slows cell growth in some situations. An inducible promoter is a promoter that directs expression of a gene where the level of expression is alterable by environmental or developmental factors such as, for example, temperature, pH, anaerobic or aerobic conditions, light, transcription factors and chemicals. Such promoters are referred to herein as "inducible" promoters, and allow one to control the timing of expression of the polypeptide.

[0071] Once expressed, the recombinant polypeptides can be purified, if desired, according to standard procedures of the art, including ammonium sulfate precipitation, affinity columns, ion exchange and/or size exclusivity chromatography, gel electrophoresis and the like (see, generally, R. Scopes, *Protein Purification*, Springer-Verlag, N.Y. (1982), Deutscher, *Methods in Enzymology Vol. 182: Guide to Protein Purification.*, Academic Press, Inc. N.Y. (1990)). Substantially pure compositions of at least about 90 to 95% homogeneity are preferred, and 98 to 99% or more homogeneity are most preferred. Once purified, partially or to homogeneity as desired, the polypeptides may then be used (*e.g.*, treatment of inflammatory diseases in pre-clinical or clinical studies).

[0072] To facilitate purification of transporters of the invention, the nucleic acids that encode the polypeptides can also include a coding sequence for an epitope or "tag" for which an affinity binding reagent is available. Examples of suitable epitopes include the myc and V-5 reporter genes; expression vectors useful for recombinant production of polypeptides having these epitopes are commercially available (*e.g.*, Invitrogen (Carlsbad CA) vectors pcDNA3.1/Myc-His and pcDNA3.1/V5-His are suitable for expression in mammalian cells; Invitrogen (Carlsbad, CA) vectors pBlueBacHis and Gibco (Gaithersburg, MD) vectors pFastBacHT are suitable for expression in insect cells). Additional expression vectors suitable for attaching a tag to the proteins of the invention, and corresponding detection systems are commercially available (*e.g.*, FLAG" (Kodak, Rochester NY). Another example of a suitable tag is a polyhistidine sequence, which is capable of binding to metal chelate affinity ligands. Typically, six adjacent histidines are used, although one can use

more or less than six. Suitable metal chelate affinity ligands that can serve as the binding moiety for a polyhistidine tag include nitrilo-tri-acetic acid (NTA) (Hochuli, E. (1990) "Purification of recombinant proteins with metal chelating adsorbents" In Genetic Engineering: Principles and Methods, J.K. Setlow, Ed., Plenum Press, NY; commercially available from Qiagen (Santa Clarita, CA)).

V. Methods of identifying agents, conjugates or conjugate moieties that are substrates of a transporter

[0073] Agents known or suspected to have pharmacological activity can be screened directly for their capacity to act as substrates of one of the transporters of the invention. Alternatively, conjugate moieties can be screened as substrates, and the conjugate moieties linked to agents having known or suspected pharmacological activity. In such methods, the conjugate moieties can be linked to an agent or other molecule during the screening process. If another molecule is used, the molecule is sometimes chosen to resemble the structure of an agent ultimately intended to be linked to the conjugate moiety for pharmaceutical use. The screening is typically performed on cells expressing a transporter. Optionally, hOst4, or other transporters of the invention, is coexpressed with a co-transporter having the amino acid sequence designated SEQ ID NO:20 (or allelic variants thereof, or variants having at least 90% sequence identity to SEQ ID NO:20 over the entire length of the co-transporter). The co-transporter is preferably encoded by a cDNA sequence designated SEQ ID NO:19 or allelic variants thereof or variants having at least 90% sequence identity thereto over the entire length of the co-transporter. The preceding disclosure with respect to fragments of transporters and nucleic acids encoding the same, and to expression of transporters applies mutatis mutandis to SEQ ID NOS:19 and 20. In some methods, the cells are transfected with DNA encoding a transporter, and optionally a co-transporter. Optionally, hOst4 or other transporter of the invention, is coexpressed with a co-transporter having the amino acid sequence designated SEQ ID NO:20. The activity of some transporters is augmented by a transporter modifying protein. Oocytes, human embryonic kidney cells (HEKs), CaCo-2, MDCK (Madin-Darby canine kidney), CHO (Chinese hamster ovary), Sf9, Sf21, High-5, or any cell line suitable for the expression of transporters may be transfected.

[0074] In other methods, natural cells express a transporter. In some methods, a transporter of the invention is the only transporter expressed. In other methods, cells express a transporter of the invention in combination with other transporters, including cotransporters of the exemplified transporters. For example, in some methods, cells

expressing at least two of hOst-1 to –4 are used. In still other methods, agents, conjugate moieties or conjugates are screened on different cells expressing different transporters. Agents, conjugate moieties or conjugates can be screened either for specificity for one transporter or for capacity to be substrates to several transporters. In some methods, agents, conjugate moieties or conjugates are screened for capacity to be a substrate for a known transporter that effects transport through an apical plasma membrane of epithelial cells lining the colon, as well as through one of the transporters of the invention. Examples of such transporters are described by WO03/065982 (e.g., MCT1, MCT4, ATBO, OCTN2, NADC1 or NADC2). Agents, conjugates, or conjugate moieties that are a substrate for both a transporter expressed in the basolateral membrane of epithelial cells and apical plasma membrane can more easily pass through the epithelial cells lining the intestine. Agents, conjugate moieties or conjugates with specificity for a particular transporter can be useful for limiting uptake to certain tissues or avoiding interaction between drugs. Agents, conjugate moieties or conjugates that are substrates for multiple transporters are useful for maximum uptake. Other types of known transporter for transport of organic solvents include the sodium coupled bile acid transporter, the sodium independent organic anion transporter, the organic anion transporter, and the organic cation transporters (see Kullack *et al.*, *Semin. Liver Dis.* 20, 273-292 (2000); Borst, *J. Natl. Cancer Inst.* 92, 1295-1302; Keppler, Semin. *Liver Dis.* 20, 265-272 (2000); Suzuki, Semin. *Liver Dis.* 20, 251-263 (2000); Saier, *Microbiol. Mol. Biol. Rev.* 64 354-411 (2000). Methods of screening agents, conjugates or conjugate moieties for passage through cells bearing a transporter are described in WO 01/20331.

[0075] Internalization of a compound evidencing passage through transporters can be detected by detecting a signal from within a cell from any of a variety of reporters. The reporter can be as simple as a label such as a fluorophore, a chromophore, a radioisotope, Confocal imaging can also be used to detect internalization of a label as it provides sufficient spatial resolution to distinguish between fluorescence on a cell surface and fluorescence within a cell; alternatively, confocal imaging can be used to track the movement of compounds over time. In another approach, internalization of a compound is detected using a reporter that is a substrate for an enzyme expressed within a cell. Once the complex is internalized, the substrate is metabolized by the enzyme and generates an optical signal or radioactive decay that is indicative of uptake. Light emission can be monitored by commercial PMT-based instruments or by CCD-based imaging systems. In addition, assay methods utilizing LCMS detection of the transported compounds or electrophysiological

signals indicative of transport activity are also employed. Measurements can also be performed in vivo by administering an agent, conjugate or conjugate moiety to an experimental animal and monitoring uptake into the plasma. The agent, conjugate or conjugate moiety can be administered orally or directly into part of the intestine in which a transporter of interest is expressed such as the small intestine or the colon.

[0076] In some methods, multiple agents, conjugates or conjugate moieties are screened simultaneously and the identity of each agent, conjugate moiety or conjugate moiety is tracked using tags linked to the agents, conjugates or conjugate moieties. In some methods, a preliminary step is performed to determine binding of an agent, conjugate or conjugate moiety to a transporter. Although not all agents, conjugates or conjugate moieties that bind to a transporter are substrates of the transporter, observation of binding is an indication that allows one to reduce the number of candidate substrates from an initial repertoire. In some methods, substrate capacity of an agent, conjugate or conjugate moiety is tested in comparison with a reference substrate of a transporter. For example, 3H-taurocholate is suitable as a reference. The comparison can either be performed in separate parallel assays in which an agent, conjugate or conjugate moiety under test and the reference substrate are compared for uptake on separate samples of the same cells. Alternatively, the comparison can be performed in a competition format in which an agent, conjugate or conjugate moiety under test and the reference substrate are applied to the same cells. Typically, the agent, conjugate or conjugate moiety and the reference substrate are differentially labeled in such assays.

[0077] In such comparative assays, the Vmax of an agent, conjugate moiety, or conjugate comprising an agent and conjugate moiety tested can be compared with that of the reference substrate (e.g., taurocholate or estrone-3-sulfate). If an agent, conjugate moiety or conjugate has a Vmax of at least 1, 5, 10. 20, 50% of the reference substrate for the transporter then the agent, conjugate moiety or conjugate can be considered to be a substrate for the transporter. In general, the higher the Vmax of the agent, conjugate moiety or conjugate relative to that of the reference substrate the better. Therefore, agents, conjugate moieties or conjugates having Vmax's of at least 50%, 100%, 150% or 200% of the Vmax of the reference substrate for the transporter are screened in some methods. The agents to which conjugate moieties are linked can by themselves show little or no detectable substrate activity for the transporter (*e.g.*, Vmax relative to that of a reference substrate of less than 0.1 or 1%).

[0078] In some methods, the Vmax of an agent, conjugate moiety or conjugate is also determined relative to the reference substrate for a second transporter. Such screening

may reveal that the agent, conjugate moiety or conjugate is a better substrate for one transporter than another. The relative capacities of a substrate for two transporters can be compared by a comparison of the ratios of Vmax of the agent, conjugate moiety or conjugate and taurocholate for the respective transporters.

VI. Agents, Conjugates and Conjugate Moieties to be Screened

[0079] Compounds constituting agents, conjugates or conjugate moieties to be screened can be naturally occurring or synthetic molecules. Natural sources include sources such as, *e.g.*, marine microorganisms, algae, plants, and fungi. Alternatively, compounds to be screened can be from combinatorial libraries of agents, including peptides or small molecules, or from existing repertories of chemical compounds synthesized in industry, *e.g.*, by the chemical, pharmaceutical, environmental, agricultural, marine, cosmeceutical, drug, and biotechnological industries. Compounds can include, *e.g.*, pharmaceuticals, therapeutics, environmental, agricultural, or industrial agents, pollutants, cosmeceuticals, drugs, organic compounds, lipids, glucocorticoids, antibiotics, peptides, sugars, carbohydrates, and chimeric molecules.

[0080] Combinatorial libraries can be produced for many types of compounds that can be synthesized in a step-by-step fashion (see *e.g.*, Ellman & Bunin, *J Amer Chem Soc*, 114:10997, 1992 (benzodiazepine template), WO 95/32184 (oxazolone and aminidine template), WO 95/30642 (dihydrobenzopyran template) and WO 95/35278 (pyrrolidine template). Libraries of compounds are usually synthesized by solid phase chemistry on particle. However, solution-phase library synthesis can also be useful. Strategies for combinatorial synthesis are described by Dolle & Nelson, *J. Combinatorial Chemistry 1.* 235-282 (1999)) (incorporated by reference in its entirety for all purposes). Synthesis is typically performed in a cyclic fashion with a different monomer or other component being added in each round of synthesis. Some methods are performed by successively fractionating an initial pool. For example, a first round of synthesis is performed on all supports. The supports are then divided into two pools and separate synthesis reactions are performed on each pool. The two pools are then further divided, each into a further two pools and so forth. Other methods employ both splitting and repooling. For example, after an initial round of synthesis, a pool of compounds is split into two for separate syntheses in a second round. Thereafter, aliquots from the separate pools are recombined for a third round of synthesis. Split and pool methods result in a pool of mixed compounds. These methods are particularly amenable for tagging as described in more detail below. The size of libraries generated by

such methods can vary from 2 different compounds to 10^4, 10^6, 10^8, or 10^{10}, or any range therebetween.

[0081] Preparation of encoded libraries is described in a variety of publications including Needels, *et al., Proc. Natl. Acad. Sci. USA* 1993, 90, 10700; Ni, *et al., J. Med. Chem.* 1996, 39, 1601, WO 95/12608, WO 93/06121, WO 94/08051, WO 95/35503 and WO 95/30642 (each of which is incorporated by reference in its entirety for all purposes). Methods for synthesizing encoded libraries typically involve a random combinatorial approach and the chemical and/or enzymatic assembly of monomer units. For example, the method typically includes steps of: (a) apportioning a plurality of solid supports among a plurality of reaction vessels; (b) coupling to the supports in each reaction vessel a first monomer and a first tag using different first monomer and tag combinations in each different reaction vessel; (c) pooling the supports; (d) apportioning the supports among a plurality of reaction vessels; (e) coupling to the first monomer a second monomer and coupling to either the solid support or to the first tag a second tag using different second monomer and second tag combinations in each different reaction vessel; and optionally repeating the coupling and apportioning steps with different tags and different monomers one to twenty or more times. The monomer set can be expanded or contracted from step to step; or the monomer set could be changed completely for the next step (*e.g.*, amino acids in one step, nucleosides in another step, carbohydrates in another step). A monomer unit for peptide synthesis, for example, can include single amino acids or larger peptide units, or both.

[0082] Compounds synthesizable by such methods include polypeptides, beta-turn mimetics, polysaccharides, phospholipids, hormones, prostaglandins, steroids, aromatic compounds, heterocyclic compounds, benzodiazepines, oligomeric N-substituted glycines and oligocarbamates. Prepared combinatorial libraries are also available from commercial sources (*e.g.*, ChemRx, South San Francisco, CA).

[0083] Some compounds to be screened are variants of known transporter substrates. Some compounds to be screened are bile salts or acids, steroids, ecosanoids, or natural toxins or analogs thereof, as described by Smith, *Am. J. Physiol.* 2230, 974-978 (1987); Smith, Am. *J. Physiol.* 252, G479-G484 (1993); Boyer, *Proc. Natl. Acad. Sci. USA* 90, 435-438 (1993); Fricker, Biochem. J. 299, 665-670 (1994); Ficker, *Biochem J.* 299, 665-670 (1994); Ballatori, *Am. J. Physiol.* 278.

VII. Linkage of Agents to Conjugate Moieties

[0084] Conjugates of this invention can be prepared by either by direct conjugation of an agent to a conjugate moiety, wherein the resulting covalent bond is cleavable *in vivo*, or by covalently coupling a difunctionalized linker precursor with an agent to a conjugate moiety. The linker precursor is selected to contain at least one reactive functionality that is complementary to at least one reactive functionality on the agent and at least one reactive functionality on the conjugate moiety. Such complementary reactive groups are well known in the art as illustrated below:

COMPLEMENTARY BINDING CHEMISTRIES

First Reactive Group	Second Reactive Group	Linkage
hydroxyl	carboxylic acid	ester
hydroxyl	haloformate	carbonate
thiol	carboxylic acid	thioester
thiol	haloformate	thiocarbonate
amine	carboxylic acid	amide
hydroxyl	isocyanate	carbamate
hydroxyl	haloformate	carbamate
amine	isocyanate	urea
carboxylic acid	carboxylic acid	anhydride
hydroxyl	phosphorus acid	phosphonate or phosphate ester

[0085] In addition to the complementary chemistry of the functional groups on the linker to both the agent and conjugate moiety, the linker (when employed) is also selected to be cleavable *in vivo*. Cleavable linkers are well known in the art and are selected such that at least one of the covalent bonds of the linker that attaches the agent to the conjugate moiety can be broken *in vivo* thereby providing for the agent or active metabolite thereof to be available to the systemic blood circulation. The linker is selected such that the reactions required to break the cleavable covalent bond are favored at the physiological site *in vivo* which permits agent (or active metabolite thereof) release into the systemic blood circulation. The selection of suitable cleavable linkers to provide effective concentrations of the agent or active metabolite thereof for release into the systemic blood circulation can be evaluated using endogenous enzymes in standard *in vitro* assays to provide a correlation to *in vivo* cleavage of the agent or active metabolite thereof from the conjugate, as is well known in the art. It is recognized that the exact cleavage mechanism employed is not critical to the methods of this invention provided, of course, that the conjugate cleaves *in vivo* in some form to provide for the agent or active metabolite thereof for sustained release into the systemic blood circulation.

[0086] In another approach, a conjugate moiety and agent are each attached to moieties having mutual affinity for each (*e.g.*, avidin or streptavidin and biotin, or hexahistidine and Ni^{2+}). In another approach, both agent and conjugate moiety are linked to a solid phase. Examples of such supports include nanoparticles (see, *e.g.*, US Pats. 5,578,325 and 5,543,158), molecular scaffolds, liposomes (see, *e.g.*, Deshmuck, D.S., *et al.*, *Life Sci.* 28:239-242 (1990), and Aramaki, Y., *et al.*, *Pharm. Res.* 10:1228-1231 (1993), protein cochleates (stable protein-phospholipid-calcium precipitates; see, *e.g.*, Chen *et al.*, *J. Contr. Rel.* 42:263-272 (1996), and clathrate complexes. These supports can be used to attach other active molecules. Certain supports such as nanoparticles can also be used to encapsulate desired compounds. An agent can be linked to a support via a cleavable linkage allowing separation of the agent after uptake through a transporter.

[0087] Examples of cleavable linkers suitable for use as described above include nucleic acids with one or more restriction sites, or peptides with protease cleavage sites (see, *e.g.*, US 5,382,513). Other exemplary linkers that can be used are available from Pierce Chemical Company in Rockford, Illinois; suitable linkers are also described in EPA 188,256; U.S. Pat. Nos. 4,671,958; 4,659,839; 4,414,148; 4,669,784; 4,680,338, 4,569, 789 and 4,589,071; and in Eggenweiler, H.M, *Drug Discovery Today, 3*: 552 (1998), each of which is incorporated in its entirety for all purposes.

[0088] There are many existing drugs for which uptake can be improved through the colon. Drugs suitable for conversion to prodrugs that are capable of uptake from the colon typically contain one or more of the following functional groups to which a promoiety may be conjugated: primary or secondary amino groups, hydroxyl groups, carboxylic acid groups, phosphonic acid groups, or phosphoric acid groups.

[0089] Examples of drugs containing carboxyl groups include, for instance, angiotensin-converting enzyme inhibitors such as alecapril, captopril, 1-[4-carboxy-2-methyl-2R,4R-pentanoyl]-2,3-dihydro-2S-indole-2-carboxylic acid, enalaprilic acid, lisinopril, N-cyclopentyl-N-[3-[(2,2-dimethyl-l-oxopropyl)thio]-2-methyl-1-oxopropyl]glycine, pivopril, quinaprilat, (2R, 4R)-2-hydroxyphenyl)-3-(3-mercaptopropionyl)-4-thiazolidinecarboxylic acid, (S) benzamido-4-oxo-6-phenylhexenoyl-2-carboxypyrrolidine, [2S-1 [R*(R*))]] 2α, 3αβ, 7αβ]-1 [2-[[1-carboxy-3-phenylpropyl]-amino]-1-oxopropyl]octahydro-lH-indole-2-carboxylic acid, [3S-1[R*(R*))]], 3R*]-2-[2-[[1-carboxy-3-phenylpropyl]-amino]-1-oxopropyl]-1,2,3,4-tetrahydro-3-isoquinolone carboxylic acid, and tiopronin; cephalosporin antibiotics such as cefaclor, cefadroxil, cefamandole, cefatrizine, cefazedone, cefazuflur,

cefazolin, cefbuperazone, cefixime, cefmenoxime, cefmetazole, cefodizime, cefonicid, cefoperazone, ceforanide, cefotaxime, cefotefan, cefotiam, cefoxitin, cefpimizole, cefpirome, cefpodoxime, cefroxadine, cefsulodin, cefpiramide, ceftazidime, ceftezole, ceftizoxime, ceftriaxone, cefuroxime, cephacetrile, cephalexin, cephaloglycin, cephaloridine, cephalosporin, cephanone, cephradine, and latamoxef; penicillins such as amoxycillin, ampicillin, apalcillin, azidocillin, azlocillin, benzylpencillin, carbenicillin, carfecillin, carindacillin, cloxacillin, cyclacillin, dicloxacillin, epicillin, flucloxacillin, hetacillin, methicillin, mezlocillin, nafcillin, oxacillin, phenethicillin, piperazillin, sulbenicllin, temocillin, and ticarcillin; thrombin inhibitors such as argatroban, melagatran, and napsagatran; influenza neuraminidase inhibitors such as zanamivir and peramivir; non-steroidal antiinflammatory agents such as acametacin, alclofenac, alminoprofen, aspirin (acetylsalicylic acid), 4-biphenylacetic acid, bucloxic acid, carprofen, cinchofen, cinmetacin, clometacin, clonixin, diclenofac, diflunisal, etodolac, fenbufen, fenclofenac, fenclosic acid, fenoprofen, ferobufen, flufenamic acid, flufenisal, flurbiprofin, fluprofen, flutiazin, ibufenac, ibuprofen, indomethacin, indoprofen, ketoprofen, ketorolac, lonazolac, loxoprofen, meclofenamic acid, mefenamic acid, 2-(8-methyl-10,11-dihydro-11-oxodibenz[b,f]oxepin-2-yl)propionic acid, naproxen, nifluminic acid, O-(carbamoylphenoxy)acetic acid, oxoprozin, pirprofen, prodolic acid, salicylic acid, salicylsalicylic acid, sulindac, suprofen, tiaprofenic acid, tolfenamic acid, tolmetin and zopemirac; prostaglandins such as ciprostene, 16-deoxy-16-hydroxy-16-vinyl prostaglandin E_2, 6,16-dimethylprostaglandin E_2, epoprostostenol, meteneprost, nileprost, prostacyclin, prostaglandins E_1, E_2, or $F_{2\alpha}$, and thromboxane A_2; quinolone antibiotics such as acrosoxacin, cinoxacin, ciprofloxacin, enoxacin, flumequine, naladixic acid, norfloxacin, ofloxacin, oxolinic acid, pefloxacin, pipemidic acid, and piromidic acid; other antibiotics such as aztreonam, imipenem, meropenem, and related carbopenem antibiotics.

[0090] Representative drugs containing amine groups include: acebutalol, albuterol, alprenolol, atenolol, bunolol, bupropion, butopamine, butoxamine, carbuterol, cartelolol, colterol, deterenol, dexpropanolol, diacetolol, dobutamine, exaprolol, exprenolol, fenoterol, fenyripol, labotolol, levobunolol, metolol, metaproterenol, metoprolol, nadolol, pamatolol, penbutalol, pindolol, pirbuterol, practolol, prenalterol, primidolol, prizidilol, procaterol, propanolol, quinterenol, rimiterol, ritodrine, solotol, soterenol, sulfiniolol, sulfinterol, sulictidil, tazaolol, terbutaline, timolol, tiprenolol, tipridil, tolamolol, thiabendazole, albendazole, albutoin, alendronate, alinidine, alizapride, amiloride, aminorex,

aprinocid, cambendazole, cimetidine, cisapride, clonidine, cyclobenzadole, delavirdine, efegatrin, etintidine, fenbendazole, fenmetazole, flubendazole, fludorex, gabapentin, icadronate, lobendazole, mebendazole, metazoline, metoclopramide, methylphenidate, mexiletine, neridronate, nocodazole, oxfendazole, oxibendazole, oxmetidine, pamidronate, parbendazole, pramipexole, prazosin, pregabalin, procainamide, ranitidine, tetrahydrazoline, tiamenidine, tinazoline, tiotidine, tocainide, tolazoline, tramazoline, xylometazoline, dimethoxyphenethylamine, N-[3(R)-[2-piperidin-4-yl)ethyl]-2-piperidone-l-yl]acetyl-3(R)-methyl-β-alanine, adrenolone, aletamine, amidephrine, amphetamine, aspartame, bamethan, betahistine, carbidopa, clorprenaline, chlortermine, dopamine, L-Dopa, ephrinephrine etryptamine, fenfluramine, methyldopamine, norepinephrine, tocainide, enviroxime, nifedipine, nimodipine, triamterene, norfloxacin, and similar compounds such as pipedemic acid, 1-ethyl-6-fluoro-1,4dihydro-4-oxo-7-(1-piperazinyl)-1, 8-napthyridine-3-carboxylic acid, 1-cyclopropyl-6-fluoro-1, and 4-dihydro-4-oxo-7-(piperazinyl)-3-quinolinecarboxylic acid.

[0091] Representative drugs containing hydroxy groups include: steroidal hormones such as allylestrenol, cingestol, dehydroepiandrosteron, dienostrol, diethylstilbestrol, dimethisteron, ethyneron, ethynodiol, estradiol, estron, ethinyl estradiol, ethisteron, lynestrenol, mestranol, methyl testosterone, norethindron, norgestrel, norvinsteron, oxogeston, quinestrol, testosterone, and tigestol; tranquilizers such as dofexazepam, hydroxyzin, lorazepam, and oxazepam; neuroleptics such as acetophenazine, carphenazine, fluphenazine, perphenyzine, and piperaetazine; cytostatics such as aclarubicin, cytarabine, decitabine, daunorubicin, dihydro-5-azacytidine, doxorubicin, epirubicin, estramustin, etoposide, fludarabine, gemcitabine, 7-hydroxychlorpromazin, nelarabine, neplanocin A, pentostatin, podophyllotoxin, tezacitabine, troxacitabine, vinblastin, vincristin, and vindesin; hormones and hormone antagonists such as buserilin, gonadoliberin, icatibrant, and leuprorelin acetate; antihistamines such as terphenadine; analgesics such as diflunisal, naproxol, paracetamol, salicylamide, and salicyclic acid; antibiotics such as azidamphenicol, azithromycin, camptothecin, cefamandol, chloramphenicol, clarithromycin, clavulanic acid, clindamycin, demeclocyclin, doxycyclin, erythromycin, gentamycin, imipenem, latamoxef, metronidazole, neomycin, novobiocin, oleandomycin, oxytetracyclin, tetracycline, thiamenicol, and tobramycin; antivirals such as acyclovir, dideoxydidehydrocytidine, dideoxycytosine, 1-(2-deoxy-2-methylene-beta-D-erythro-pentofuranosyl)cytidine, fluoro-dideoxydidehydrocytidine, fluorodideoxycytosine, FMAU (1-(2-deoxy-2-fluoro-beta-D-

arabinofuranosyl)thymine), deoxy-5-fluoro-3'-thiacytidine, 2'-fluoro-ara-dideoxyinosine, ganciclovir, lamivudine, penciclovir, SddC, stavudine, 5-trifluoromethyl-2'-deoxyuridine, zalcitabine, and zidovudine; bisphosphonates such as EB-1053 (1-hydroxy-3-(1-pyrrolidinyl)propylidene-1,1-bisphosphonate), etidronate, ibandronate, olpadronate, residronate, 1-hydroxy-2-(imidazo [1,2-a] pyridin-3-yl) ethylidene]-bisphosphonic acid, and zolendronate; protease inhibitors such as ciprokiren, enalkiren, ritonavir, saquinavir, and terlakiren; prostaglandins such as arbaprostil, carboprost, misoprostil, and prostacydin; antidepressives such as 8-hydroxychlorimipramine and 2-hydroxyimipramine; antihypertonics such as sotarol and fenoldopam; anticholinerogenics such as biperidine, procyclidin and trihexyphenidal; antiallergenics such as cromolyn; glucocorticoids such as betamethasone, budenosid, chlorprednison, clobetasol, clobetasone, corticosteron, cortisone, cortodexon, dexamethason, flucortolon, fludrocortisone, flumethasone, flunisolid, fluprednisolon, flurandrenolide, flurandrenolon acetonide, hydrocortisone, meprednisone, methylpresnisolon, paramethasone, prednisolon, prednisol, triamcinolon, and triamcinolon acetonide; narcotic agonists and antagonists such as apomorphine, buprenorphine, butorphanol, codein, cyclazocin, hydromorphon, ketobemidon, levallorphan, levorphanol, metazocin, morphine, nalbuphin, nalmefen, naloxon, nalorphine, naltrexon, oxycodon, oxymorphon, and pentazocin; stimulants such asmazindol and pseudoephidrine; anaesthetics such as hydroxydion and propofol; β-receptor blockers such as acebutolol, albuterol, alprenolol, atenolol, betazolol, bucindolol, cartelolol, celiprolol, cetamolol, labetalol, levobunelol, metoprolol, metipranolol, nadolol, oxyprenolol, pindolol, propanolol, and timolol; α-sympathomimetics such as adrenalin, metaraminol, midodrin, norfenefrin, octapamine, oxedrin, oxilofrin, oximetazolin, and phenylefrin; β-sympathomimetics such as bamethan, clenbuterol, fenoterol, hexoprenalin, isoprenalin, isoxsuprin, orciprenalin, reproterol, salbutamol, and terbutalin; bronchodilators such as carbuterol, dyphillin, etophyllin, fenoterol, pirbuterol, rimiterol and terbutalin; cardiotonics such as digitoxin, dobutamin, etilefrin, and prenalterol; antimycotics such as amphotericin B, chlorphenesin, nystatin, and perimycin; anticoagulants such as acenocoumarol, dicoumarol, phenprocoumon, and warfarin; vasodilators such as bamethan, dipyrimadol, diprophyllin, isoxsuprin, vincamin and xantinol nicotinate; antihypocholesteremics such as compactin, eptastatin, mevinolin, and simvastatin; miscellaneous drugs such as bromperidol (antipsychotic), dithranol (psoriasis) ergotamine (migraine) ivermectin (antihelminthic), metronidazole and secnizadole

(antiprotozoals), nandrolon (anabolic), propafenon and quinadine (antiarythmics), quetiapine (CNS), serotonin (neurotransmitter), and silybin (hepatic disturbance).

[0092] Representative drugs containing phosphonic acid moieties include: adefovir, alendronate, (N6-[2-methylthio)ethyl]-2-[3,3,3-trifluoropropylthio]-5'-adenylic acid, BMS-187745 (a squalene synthase inhibitor from Bristol-Meyers Squibb Inc.), ceronapril, CGP-24592 (Novartis, Inc.), DL-(E)-2-amino-4-methyl-5-phosphono-3-pentenoic acid; 4-methyl-APPA, CGP-39551 (ethyl esters of (DL-[E]-2-amino-4-methyl-5-phosphono-3-pentenoic acid)), CGP-40116 (a competitive NMDA antagonist by Novartis Inc.), cidofovir, clodronate, EB-1053 (1-hydroxy-3-(1-pyrrolidinyl)propylidene-1,1-bisphosphonate), etidronate, fanapanel, foscarnet, fosfomycin, fosinopril, fosinoprilat, ibandronate, midafotel, neridronate, olpadronate, pamidronate, residronate, tenofovir, tiludronate, [2-(8,9-dioxo-2,6-diazabicyclo[5.2.0]non-1(7)-en-2-yl)ethyl]phosphonic acid, 1-hydroxy-2-(imidazo [1,2-a] pyridin-3-yl) ethylidene]-bisphosphonic acid, and zolendronate.

[0093] Representative drugs containing phosphoric acid moieties include: bucladesine, choline alfoscerate, citocoline, fludarabine phosphate, fosopamine, GP-668, perifosine, triciribine phosphate, and phosphate derivatives of nucleoside analogs which require phophorylation for activity, such as lamivudine, acyclovir, azidothymidine, E-5-(2-bromovinyl)-2'-deoxyuridine, dideoxycytosine, dideoxyinosine, FMAU (1-(2-deoxy-2-fluoro-beta-D-arabinofuranosyl)thymine), deoxy-5-fluoro-3'-thiacytidine, ganciclovir, gemcitabine, (R)-9-[4-Hydroxy-2-(hydroxymethy)butyl]guanine, lamivudine, penciclovir and the like.

[0094] Preferred drugs for modification to prodrugs capable of intestional absorption and incorporation into sustained release formulations include the following compounds: analgesics and/or antiinflammatory agents selected from the group consisting of acetaminophen, buprenorphine, diclofenac, diflunisal, fenoprofen, ibuprofen, indomethacin, ketoprofen, mefenamic acid, meptazinol, morphine, oxycodone, pentazocine, pethidine, tolmetin, and tramadol; antihypertensive agents selected from the group consisting of captopril, diltiazem, methyldopa, metoprolol, prazosin, propranolol, quinapril, sotalol, and timolol; antibiotic agents selected from the group consisting of amoxicillin, ampicillin, aztreonam, cefaclor, cefadroxil, cefixime, cefotaxime, cefoxitin, cefpodoxime, ceftizoxime, ceftriaxone, cefuroxime, cephalexin, ciproflaxacin, clindamycin, erythromycin, imipenem, mandol, meropenem, metronidazole, and tobramycin; antiviral agents selected from the group consisting of acyclovir, delavirdine, didanosine, foscarnet, ganciclovir, indinavir, lamivudine, nelfinavir, penciclovir, ritonavir, saquinavir, stavudine, zalcitabine, and zidovudine;

bronchodilator and or anti-asthmatic agents selected from the group consisting of salbutamol and terbutaline; antiarrhythmic agents selected from the group consisting of mexiletine, procainamide, and tocainide; centrally acting substances selected from the group consisting of baclofen, benserazide, bupropion, carbidopa, gabapentin, levodopa, methylphenildate, pramipexole, pregabalin, quetiapine, ropinirole, and vigabatrin; cytostatics and metastasis inhibitors selected from the group consisting of cytarabine, decitabine, docetaxal, flutamide, gemcitabine, paclitaxel, and pentostatin; and, agents for treatment of gastrointestinal disorders selected from the group consisting of cisapride, metoclopramide, and misoprostol.

VIII. Pharmaceutical Compositions and Methods of Treatment

[0095] Agents that are themselves substrates for a transporter or which are linked to conjugate moieties that are substrates for a transporter can be can be incorporated into pharmaceutical compositions. Usually, although not necessarily, such pharmaceutical compositions are designed for oral administration. Oral administration of such compositions results in uptake through the intestine via a transporter and entry into the systemic circulation. The pharmaceutical composition can thus be efficiently delivered to a wide range of tissues in the body.

[0096] Agents optionally linked to a conjugate moiety are combined with pharmaceutically-acceptable, non-toxic carriers of diluents, which are defined as vehicles commonly used to formulate pharmaceutical compositions for animal or human administration. The diluent is selected so as not to affect the biological activity of the combination. Examples of such diluents are distilled water, buffered water, physiological saline, PBS, Ringer's solution, dextrose solution, and Hank's solution. In addition, the pharmaceutical composition or formulation can also include other carriers, adjuvants, or non-toxic, nontherapeutic, nonimmunogenic stabilizers, excipients and the like. The compositions can also include additional substances to approximate physiological conditions, such as pH adjusting and buffering agents, toxicity adjusting agents, wetting agents, detergents and the like (see, *e.g.*, Remington's Pharmaceutical Sciences, Mace Publishing Company, Philadelphia, PA, 17th ed. (1985); for a brief review of methods for drug delivery, see, Langer, *Science* 249:1527-1533 (1990); each of these references is incorporated by reference in its entirety).

[0097] Pharmaceutical compositions for oral administration can be in the form of *e.g.*, tablets, pills, powders, lozenges, sachets, cachets, elixirs, suspensions, emulsions, solutions, or syrups. Some examples of suitable excipients include lactose, dextrose, sucrose,

sorbitol, mannitol, starches, gum acacia, calcium phosphate, alginates, tragacanth, gelatin, calcium silicate, microcrystalline cellulose, polyvinylpyrrolidone, cellulose, sterile water, syrup, and methyl cellulose. Preserving agents such as methyl- and propylhydroxy-benzoates; sweetening agents; and flavoring agents can also be included. Depending on the formulation, compositions can provide quick, sustained or delayed release of the active ingredient after administration to the patient. The tablets or pills of the present invention may be coated or otherwise compounded to provide a dosage form affording the advantage of prolonged action. For example, the tablet or pill can comprise an inner dosage and an outer dosage component, the latter being in the form of an envelope over the former. The two components can be separated by an enteric layer which serves to resist disintegration in the stomach and permit the inner component to pass intact into the duodenum or to be delayed in release. A variety of materials can be used for such enteric layers or coatings, such materials including a number of polymeric acids and mixtures of polymeric acids with such materials as shellac, cetyl alcohol, and cellulose acetate.

[0098] For preparing solid compositions such as tablets, the principal active ingredient is mixed with a pharmaceutical excipient to form a solid preformulation composition containing a homogeneous mixture of a compound of the present invention. When referring to these preformulation compositions as homogeneous, it is meant that the active ingredient is dispersed evenly throughout the composition so that the composition may be readily subdivided into equally effective unit dosage forms such as tablets, pills and capsules. This solid preformulation is then subdivided into unit dosage forms of the type described above containing from, for example, 0.1 mg to about 2 g of the active agent.

[0099] The compositions can be administered for prophylactic and/or therapeutic treatments. A therapeutic amount is an amount sufficient to remedy a disease state or symptoms, or otherwise prevent, hinder, retard, or reverse the progression of disease or any other undesirable symptoms in any way whatsoever. In prophylactic applications, compositions are administered to a patient susceptible to or otherwise at risk of a particular disease or infection. Hence, a "prophylactically effective" is an amount sufficient to prevent, hinder or retard a disease state or its symptoms. In either instance, the precise amount of compound contained in the composition depends on the patient's state of health and weight.

[0100] An appropriate dosage of the pharmaceutical composition is readily determined according to any one of several well-established protocols. For example, animal studies (*e.g.*, mice, rats) are commonly used to determine the maximal tolerable dose of the bioactive agent per kilogram of weight. In general, at least one of the animal species tested is

mammalian. The results from the animal studies can be extrapolated to determine doses for use in other species, such as humans for example.

[0101] The pharmaceutical compositions can be administered in a variety of different ways. Examples include administering a composition containing a pharmaceutically acceptable carrier via oral, intranasal, rectal, topical, intraperitoneal, intravenous, intramuscular, subcutaneous, subdermal, transdermal, intrathecal, and intracranial methods. The route of administration depends in part on the chemical composition of the active compound and any carriers.

[0102] The components of pharmaceutical compositions are preferably of high purity and are substantially free of potentially harmful contaminants (*e.g.*, at least National Food (NF) grade, generally at least analytical grade, and more typically at least pharmaceutical grade). To the extent that a given compound must be synthesized prior to use, the resulting product is typically substantially free of any potentially toxic agents, particularly any endotoxins, which may be present during the synthesis or purification process. Compositions for parental administration are also sterile, substantially isotonic and made under GMP conditions. Compositions for oral administration need not be sterile or substantially isotonic but are usually made under GMP conditions.

IX. Other Applications of Transporters

1. Antibodies

[0103] The transporters of the invention can be used to generate antibodies. The antibodies can be polyclonal antibodies, distinct monoclonal antibodies or pooled monoclonal antibodies with different epitopic specificities. Monoclonal antibodies are made from antigen-containing fragments of the protein by standard procedures according to the type of antibody (see, *e.g.*, Kohler, *et al.*, *Nature*, 256:495, (1975); and Harlow & Lane, *Antibodies, A Laboratory Manual* (C.S.H.P., NY, 1988) Queen *et al.*, *Proc. Natl. Acad. Sci. USA* 86:10029-10033 (1989) and WO 90/07861; Dower *et al.*, WO 91/17271 and McCafferty *et al.*, WO 92/01047 (each of which is incorporated by reference for all purposes). Nonhuman antibodies are typically made by immunizing a nonhuman animal, such as a mouse, harvesting B-cells from the animal, immortalizing the cells to produce hybridomas, and selecting a hybridoma secreting an antibody having the desired binding characterisitics of a tranporter. The antibodies of the invention can be chimeric, humanized, human, mouse or other species. Phage display technology can also be used to mutagenize CDR regions of

antibodies previously shown to have affinity for the peptides of the present invention. Preferably such antibodies do not specifically bind to the skate Ost transporter. The antibodies can be further purified, for example, by binding to and elution from a support to which the polypeptide or a peptide to which the antibodies were raised is bound.

[0104] Antibodies of the invention are useful, for example, in screening cDNA expression libraries and for identifying clones containing cDNA inserts which encode structurally-related, immunocrossreactive proteins. See, for example, Aruffo & Seed, *Proc. Natl. Acad. Sci. USA* 84:8573-8577 (1977) (incorporated herein by reference in its entirety for all purposes). Antibodies are also useful to identify and/or purify immunocrossreactive proteins that are structurally related to SEQ ID NOS: 2, 4, 6, or 8 to fragments thereof used to generate the antibody.

2. Expression Monitoring Arrays

[0105] Nucleic acids encoding the transporters of the invention are also useful for inclusion on a GeneChipTM array or the like for use in expression monitoring (see US 6,040,138, . EP 853, 679 and WO97/27317). Such arrays typically contain oligonucleotide or cDNA probes to allow detection of large numbers of mRNAs within a mixture. Many of the nucleic acids included in such arrays are from genes or ESTs that have not been well characterized. Such arrays are often used to compare expression profiles between different tissues or between different conditions of the same tissue (healthy vs. diseased or drug-treated vs. control) to identify differentially expressed transcripts. The differentially expressed transcripts are then useful *e.g.*, for diagnosis of disease states, or to characterize responses of drugs. The nucleic acids of the invention can be included on GeneChipTM arrays or the like together with probes containing a variety of other genes. The present nucleic acids are particularly useful for inclusion in GeneChipTM arrays for analyzing the transport capacity of a cell and effects of drugs on the same. Nucleic acids encoding the transporters of the invention can be combined with nucleic acids encoding other transporters molecules.

3. Diagnosis

[0106] cDNA or genomic DNA of hOst-1, -2, -3 and 4, can be sequenced and compared with the exemplary sequences of SEQ. ID NOS: 1, 3, 5, and 7. Variation from one of these sequences, particularly a nucleic acid substitution giving rise to a nonconservative amino acid substitution can be indicative of disease. To perform such analysis, the presence or absence of one or more polymorphic forms *(i.e.,* a polymorphic set) is determined for a set

of the individuals, some of whom exhibit a particular trait, and some of which exhibit lack of the trait. The alleles of each polymorphism of the set are then reviewed to determine whether the presence or absence of a particular allele is associated with the trait of interest. Correlation can be performed by standard statistical methods such as a chi-squared test and statistically significant correlations between polymorphic form(s) and phenotypic characteristics are noted. Polymorphisms found to correlate with the disease trait can then be used as a basis for a genetic trait.

4. Screening for agents that agonize or antagonize transporter function

[0107] The transporters of the invention are useful for screening for agents that agonize or antagonize transporter function. Such agents can be useful as drugs in compensating for genetic variations affecting transporter functions. For example, an agonist of transporter function is useful in a patient having a genetic variation that decreases endogenous transport function, and an antagonist is useful in a patient having a genetic variation that increases endogenous transport function. An agonist is also useful in patients having normal or defective transporter function to increase uptake of another drug against a different target. The transporters can also be used to screen known drugs against targets other than transporters to determine whether the drug has an incidental effect on transport activity. A drug that has an incidental agonist activity should generally be avoided in patients having atypically high levels of transporter activity, and an antagonist should generally be avoided in patients having an atypically low level of endogenous transport activity.

[0108] Agents can be screened in cells transfected with nucleic acids encoding a transporter of the invention or in transgenic animals expressing a transporter of the invention as a transgene. Activity of an agent is monitored from its effect on transport of a known substrate such as taurocholate. Agents for screening can be obtained by producing and screening large combinatorial libraries, as described above, or can be known drugs.

[0109] Although the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. The above examples are provided to illustrate the invention, but not to limit its scope; other variants of the invention will be readily apparent to those of ordinary skill in the and are encompassed by the claims of the invention. The scope of the invention should, therefore, be determined not with reference to the above description, but instead should be determined with reference to the appended claims along with their full

scope of equivalents. All publications, references, and patent documents cited in this application are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication or patent document were so individually denoted.

WHAT IS CLAIMED IS:

1. A method of screening whether a conjugate or conjugate moiety is a substrate of a transporter, comprising:

providing a cell expressing a nucleic acid to produce a transporter in an plasma membrane of the cell, the transporter having an amino acid sequence at least 90 % identical to SEQ ID NO: 2 over the entire length of SEQ ID NO:2, wherein the transporter is coexpressed with a co-transporter having at least 90% sequence identity to SEQ ID NO:20 over the entire length of the SEQ ID NO:20;

contacting the cell with a conjugate moiety or conjugate; and

determining whether the conjugate moiety or conjugate passes through the plasma membrane via the transporter.

2. The method of claim 1, wherein transporter has the sequence of SEQ. ID NO: 2.

3. The method of claim1, wherein the co-transporter has the sequence of SEQ ID NO:20.

4. The method of claim 1, wherein the cell is a Chinese hamster ovary cell, a human embryonic kidney cell or an oocyte.

5. The method of claim 1, further comprising formuating the conjugate with a carrier as a pharmaceutical composition.

6. A method of screening agents, conjugate moieties or conjugates as substrate for intestinal transport, comprising

screening an agent, conjugate moiety or conjugate for capacity to be a substrate for a known transporter that effects transport through an apical plasma membrane of epithelial cells lining the intestine,

screening the agent, conjugate moiety or conjugate for capacity to be a substrate for a transporter having the amino acid sequence of SEQ ID NO:2.

7. The method of claim 6, further comprising formuating the conjugate with a carrier as a pharmaceutical composition.

8. The method of claim 6, wherein the known transporter effects transport through an apical plasma membrane of epithelial cells lining the intestine.

9. A method of screening agents, conjugates or conjugate moieties for capacity to agonize or antagonize a transporter, comprising

contacting a cell expressing a transporter having an amino acid sequence at least 90 % identical to SEQ ID NO: 2 over the entire length of SEQ ID NO:2, wherein the transporter is coexpressed with a co-transporter having at least 90% sequence identity to SEQ ID NO:20 over the entire length of the SEQ ID NO:20, the transporter being situated in the plasma membrane of the cell; with an agent, conjugate or conjugate moiety and a known substrate of the transporter;

determining whether the agent agonizes or antagonizes uptake of the known substrate into the cell in comparison with a control cell expressing the transporter contacted with known substrate without the agent, conjugate or conjugate moiety.

10. The method of claim 9, wherein the known substrate is taurocholate or estrone-3-sulfate.

11. The method of claim 9, wherein transporter has the sequence of SEQ. ID NO: 2.

12. The method of claim 9, wherein the co-transporter has the sequence of SEQ ID NO:20.

13. A conjugate comprising an agent linked to a conjugate moiety for a transporter having the amino acid sequence SEQ ID NO: 2, wherein the transporter is coexpressed with a co-transporter having the amino acid sequence SEQ ID NO:20, wherein the conjugate shows a Vmax of at least 1% of taurocholate for the transporter, wherein the agent has a pharmaceutical activity without the conjugate moiety, and the conjugate has a greater Vmax for the transporter than the agent without the conjugate moiety.

14. A method of manufacturing a pharmaceutical composition, comprising;

linking an agent to a conjugate moiety to form a conjugate wherein the conjugate is transported by a transporter having the amino acid sequence SEQ ID NO: 2 over the entire length of SEQ ID NO:2, wherein the transporter is coexpressed with a co-transporter having the amino acid sequence SEQ ID NO:20, with a higher Vmax than the agent alone;

formulating the conjugate with a carrier as a pharmaceutical composition.

15. The use of a conjugate comprising an agent linked to a conjugate moiety wherein the conjugate is transported by a transporter having the amino acid sequence SEQ ID NO: 2 over the entire length of SEQ ID NO:2, wherein the transporter is coexpressed with a co-transporter having the amino acid sequence SEQ ID NO:20, with a higher Vmax than the agent alone in the manufacture of a medicament.

16. The use of claim 15, wherein the medicament is a medicament for oral administration.

17. The use of claim 15, wherein the medicament is a medicament for intravenous administration.

18. Use of a drug covalently linked to a conjugate moiety through linkage cleavable in vivo to form a conjugate in the manufacture of a medicament for oral administration, wherein the conjugate is transported by a transporter having an amino acid sequence of SEQ ID NO: 2, wherein the transporter is coexpressed with a co-transporter having an amino acid sequence of SEQ ID NO:20, the transporter being situated in the plasma membrane of epithelial cells lining the intestinal lumen, wherein the conjugate has a Vmax of at least 1% of taurocholate for the transporter, wherein oral bioability of the agent due to uptake of the conjugate is improved relative to the oral bioavailabilityof the agent administered without the conjugate moiety.

19. The use of claim 18, wherein the medicament is suitable for oral administration to a human.

20. The use of claim 18, , wherein the medicament is in a sustained release oral dosage form.

SEQUENCE LISTING

<110> Zerangue, Noa

Paddon, Chris

<120> Human Organic Solute Transporters

<130> 019282-001511PC

<140> To Be Assigned

<141> 2004-01-24

<150> US 10/763,644

<151> 2004-01-22

<160> 20

<170> PatentIn version 3.1

<210> 1

<211> 1023

<212> DNA

<213> Homo sapiens

<400> 1

atggagccgg gcaggaccca gataaagctt gaccccaggt acacagcaga tcttctggag 60

gtgctgaaga ccaattacgg catcccctcc gcctgcttct ctcagcctcc cacagcagcc 120

caactcctga gagccctggg ccctgtggaa cttgccctca ctagcatcct gaccttgctg 180

gcgctgggct ccattgccat cttcctggag gatgccgtct acctgtacaa gaacaccctt 240

tgccccatca agaggcggac tctgctctgg aagagctcgg cacccacggt ggtgtctgtg 300

ctgtgctgct ttggtctctg gatccctcgt tccctggtgc tggtggaaat gaccatcacc 360

tcgttttatg ccgtgtgctt ttacctgctg atgctggtca tggtggaagg ctttgggggg 420

aaggaggcag tgctgaggac gctgagggac accccgatga tggtccacac aggcccctgc 480

tgctgctgct gcccctgctg tccacggctg ctgctcacca ggaagaagct tcagctgctg 540

atgttgggcc ctttccaata cgccttcttg aagataacgc tgaccctggt gggcctgttt 600

ctcgtccccg acggcatcta tgacccagca gacatttctg aggggagcac agctctatgg 660

atcaacactt tccttggcgt gtccacactg ctggctctct ggaccctggg catcatttcc 720

cgtcaagcca ggctacacct gggtgagcag aacatgggag ccaaatttgc tctgttccag 780

gttctcctca tcctgactgc cctacagccc tccatcttct cagtcttggc caacggtggg 840

cagattgctt gttcgcctcc ctattcctct aaaaccaggt ctcaagtgat gaattgccac 900

ctcctcatac tggagacttt tctaatgact gtgctgacac gaatgtacta ccgaaggaaa 960

gaccacaagg ttgggtatga aactttctct tctccagacc tggacttgaa cctcaaagcc 1020

taa 1023

<210> 2

<211> 340

<212> PRT

<213> Homo sapiens

<400> 2

Met Glu Pro Gly Arg Thr Gln Ile Lys Leu Asp Pro Arg Tyr Thr Ala
1 5 10 15

Asp Leu Leu Glu Val Leu Lys Thr Asn Tyr Gly Ile Pro Ser Ala Cys
20 25 30

Phe Ser Gln Pro Pro Thr Ala Ala Gln Leu Leu Arg Ala Leu Gly Pro
35 40 45

Val Glu Leu Ala Leu Thr Ser Ile Leu Thr Leu Leu Ala Leu Gly Ser
50 55 60

Ile Ala Ile Phe Leu Glu Asp Ala Val Tyr Leu Tyr Lys Asn Thr Leu
65 70 75 80

Cys Pro Ile Lys Arg Arg Thr Leu Leu Trp Lys Ser Ser Ala Pro Thr
85 90 95

```
Val Val Ser Val Leu Cys Cys Phe Gly Leu Trp Ile Pro Arg Ser Leu
            100                 105                 110

Val Leu Val Glu Met Thr Ile Thr Ser Phe Tyr Ala Val Cys Phe Tyr
        115                 120                 125

Leu Leu Met Leu Val Met Val Glu Gly Phe Gly Gly Lys Glu Ala Val
    130                 135                 140

Leu Arg Thr Leu Arg Asp Thr Pro Met Met Val His Thr Gly Pro Cys
145                 150                 155                 160

Cys Cys Cys Cys Pro Cys Cys Pro Arg Leu Leu Leu Thr Arg Lys Lys
                165                 170                 175

Leu Gln Leu Leu Met Leu Gly Pro Phe Gln Tyr Ala Phe Leu Lys Ile
            180                 185                 190

Thr Leu Thr Leu Val Gly Leu Phe Leu Val Pro Asp Gly Ile Tyr Asp
        195                 200                 205

Pro Ala Asp Ile Ser Glu Gly Ser Thr Ala Leu Trp Ile Asn Thr Phe
    210                 215                 220

Leu Gly Val Ser Thr Leu Leu Ala Leu Trp Thr Leu Gly Ile Ile Ser
225                 230                 235                 240

Arg Gln Ala Arg Leu His Leu Gly Glu Gln Asn Met Gly Ala Lys Phe
                245                 250                 255

Ala Leu Phe Gln Val Leu Leu Ile Leu Thr Ala Leu Gln Pro Ser Ile
            260                 265                 270

Phe Ser Val Leu Ala Asn Gly Gly Gln Ile Ala Cys Ser Pro Pro Tyr
        275                 280                 285

Ser Ser Lys Thr Arg Ser Gln Val Met Asn Cys His Leu Leu Ile Leu
    290                 295                 300

Glu Thr Phe Leu Met Thr Val Leu Thr Arg Met Tyr Tyr Arg Arg Lys
305                 310                 315                 320

Asp His Lys Val Gly Tyr Glu Thr Phe Ser Ser Pro Asp Leu Asp Leu
                325                 330                 335
```

Asn Leu Lys Ala
340

<210> 3

<211> 1122

<212> DNA

<213> Homo sapiens

<400> 3
atggagcagc ctgtgttcct gatgacaact gccgctcagg ccatctctgg cttcttcgtg 60

tggacggccc tgctcatcac atgccaccag atctacatgc acctgcgctg ctacagctgc 120

cccaacgagc agcgctacat cgtgcgcatc ctcttcatcg tgcccatcta cgcctttgac 180

tcctggctca gcctcctctt cttcaccaac gaccagtact acgtgtactt cggcaccgtc 240

cgcgactgct atgaggcctt ggtcatctat aatttcctga gcctgtgcta tgagtaccta 300

ggaggagaaa gttccatcat gtcggagatc agaggaaaac ccattgagtc cagctgtatg 360

tatggcacct gctgcctctg gggaaagact tattccatcg gatttctgag gttctgcaaa 420

caggccaccc tgcagttctg tgtggtgaag ccactcatgg cggtcagcac tgtggtcctc 480

caggccttcg gcaagtaccg ggatggggac tttgacgtca ccagtggcta cctctacgtg 540

accatcatct acaacatctc cgtcagcctg gccctctacg ccctcttcct cttctacttc 600

gccacccggg agctgctcag cccctacagc cccgtcctca agttcttcat ggtcaagtcc 660

gtcatctttc tttccttctg gcaaggcatg ctcctggcca tcctggagaa gtgtggggcc 720

atccccaaaa tccactcggc ccgcgtgtcg gtgggcgagg gcaccgtggc tgccggctac 780

caggacttca tcatctgtgt ggagatgttc tttgcagccc tggccctgcg gcacgccttc 840

acctacaagg tctatgctga caagaggctg gacgcacaag gccgctgtgc ccccatgaag 900

agcatctcca gcagcctcaa ggagaccatg aacccgcacg acatcgtgca ggacgccatc 960

cacaacttct cacctgccta ccagcagtac acgcagcagt ccaccctgga gcctgggccc 1020

acctggcgtg gtggcgccca cggcctctcc cgctcccaca gcctcagtgg cgcccgcgac 1080

aacgagaaga ctctcctgct cagctctgat gatgaattct ag 1122

<210> 4

<211> 373

<212> PRT

<213> Homo sapiens

<400> 4

```
Met Glu Gln Pro Val Phe Leu Met Thr Thr Ala Ala Gln Ala Ile Ser
1               5                  10                  15

Gly Phe Phe Val Trp Thr Ala Leu Leu Ile Thr Cys His Gln Ile Tyr
            20                  25                  30

Met His Leu Arg Cys Tyr Ser Cys Pro Asn Glu Gln Arg Tyr Ile Val
        35                  40                  45

Arg Ile Leu Phe Ile Val Pro Ile Tyr Ala Phe Asp Ser Trp Leu Ser
    50                  55                  60

Leu Leu Phe Phe Thr Asn Asp Gln Tyr Tyr Val Tyr Phe Gly Thr Val
65                  70                  75                  80

Arg Asp Cys Tyr Glu Ala Leu Val Ile Tyr Asn Phe Leu Ser Leu Cys
                85                  90                  95

Tyr Glu Tyr Leu Gly Gly Glu Ser Ser Ile Met Ser Glu Ile Arg Gly
            100                 105                 110

Lys Pro Ile Glu Ser Ser Cys Met Tyr Gly Thr Cys Cys Leu Trp Gly
        115                 120                 125

Lys Thr Tyr Ser Ile Gly Phe Leu Arg Phe Cys Lys Gln Ala Thr Leu
    130                 135                 140

Gln Phe Cys Val Val Lys Pro Leu Met Ala Val Ser Thr Val Val Leu
145                 150                 155                 160

Gln Ala Phe Gly Lys Tyr Arg Asp Gly Asp Phe Asp Val Thr Ser Gly
                165                 170                 175

Tyr Leu Tyr Val Thr Ile Ile Tyr Asn Ile Ser Val Ser Leu Ala Leu
            180                 185                 190

Tyr Ala Leu Phe Leu Phe Tyr Phe Ala Thr Arg Glu Leu Leu Ser Pro
        195                 200                 205

Tyr Ser Pro Val Leu Lys Phe Phe Met Val Lys Ser Val Ile Phe Leu
    210                 215                 220
```

```
Ser Phe Trp Gln Gly Met Leu Leu Ala Ile Leu Glu Lys Cys Gly Ala
225                 230                 235                 240

Ile Pro Lys Ile His Ser Ala Arg Val Ser Val Gly Glu Gly Thr Val
                245                 250                 255

Ala Ala Gly Tyr Gln Asp Phe Ile Ile Cys Val Glu Met Phe Phe Ala
            260                 265                 270

Ala Leu Ala Leu Arg His Ala Phe Thr Tyr Lys Val Tyr Ala Asp Lys
        275                 280                 285

Arg Leu Asp Ala Gln Gly Arg Cys Ala Pro Met Lys Ser Ile Ser Ser
    290                 295                 300

Ser Leu Lys Glu Thr Met Asn Pro His Asp Ile Val Gln Asp Ala Ile
305                 310                 315                 320

His Asn Phe Ser Pro Ala Tyr Gln Gln Tyr Thr Gln Gln Ser Thr Leu
                325                 330                 335

Glu Pro Gly Pro Thr Trp Arg Gly Gly Ala His Gly Leu Ser Arg Ser
            340                 345                 350

His Ser Leu Ser Gly Ala Arg Asp Asn Glu Lys Thr Leu Leu Leu Ser
        355                 360                 365

Ser Asp Asp Glu Phe
    370
```

<210> 5

<211> 1242

<212> DNA

<213> *Homo sapiens*

<400> 5

```
atgagtaatg tctcagggat cctggagaca gccggcgtcc ccctggtgtc agcgaactgg     60

ccgcagccca gccccccacc ggctgtgcca gctgggccgc agatggacca catggggaac    120

agctcccagg gggccccctg gctcttcctc acctccgcac tggcccgagg cgtctcgggg    180

atcttcgtgt ggactgccct ggtgctcacc tgccaccaga tctatctgca cctgcgctcc    240
```

tacaccgtgc cacaggagca acgttacatc atccgcctgc tcctcatcgt gcccatctac 300

gccttcgact cctggctcag cctcctcctc ctcggagacc accagtacta cgtctacttc 360

gactctgtgc gggactgcta cgaagccttt gtcatttaca gcttcctgag cctgtgtttc 420

cagtacctgg gaggcgaggg cgccatcatg gctgagattc gtggaaagcc catcaagtcc 480

agctgcttgt acggcacctg ctgcctccgg ggcatgacct actccatcgg gttcctgcgc 540

ttctgtaagc aggccactct gcagttctgc ctggtgaagc ccgtcatggc cgtcaccacc 600

atcatcctcc aggcatttgg caaataccac gacggggact tcaatgtccg cagcggctac 660

ctctatgtga ccctcatcta caacgcctcc gtcagcctcg ccctctacgc cctgttcctc 720

ttctacttca ccaccaggga gctcctgcgg cccttccagc ccgtcctcaa gttcctcacc 780

atcaaagccg tcatcttcct gtcgttctgg caagggctgc tgctggccat cctggagcgg 840

tgcggggtca tcccggaggt ggagaccagc ggcgggaaca agctgggggc tggcacgctg 900

gccgccggct accagaactt catcatctgc gtggagatgc tgttcgcctc cgtggccctg 960

cgttatgcct tcccctgcca ggtgtacgca gagaagaagg agaattcacc agccccccg 1020

gcacccatgc agagcatctc cagcggcatc agggagacag tgagccccca ggacatcgtg 1080

caggacgcca tccacaactt ctccccgcc taccagcact acacgcagca ggccacgcac 1140

gaggcgccca ggcccggcac ccaccccagc ggcggctccg gcgggagcag gaagagccgg 1200

agcctggaga agcggatgct gatcccctcg gaggacctgt ag 1242

<210> 6

<211> 413

<212> PRT

<213> Homo sapiens

<400> 6

Met Ser Asn Val Ser Gly Ile Leu Glu Thr Ala Gly Val Pro Leu Val
1 5 10 15

Ser Ala Asn Trp Pro Gln Pro Ser Pro Pro Pro Ala Val Pro Ala Gly
20 25 30

Pro Gln Met Asp His Met Gly Asn Ser Ser Gln Gly Ala Pro Trp Leu
35 40 45

```
Phe Leu Thr Ser Ala Leu Ala Arg Gly Val Ser Gly Ile Phe Val Trp
    50                  55                  60

Thr Ala Leu Val Leu Thr Cys His Gln Ile Tyr Leu His Leu Arg Ser
65                  70                  75                  80

Tyr Thr Val Pro Gln Glu Gln Arg Tyr Ile Ile Arg Leu Leu Leu Ile
                85                  90                  95

Val Pro Ile Tyr Ala Phe Asp Ser Trp Leu Ser Leu Leu Leu Leu Gly
            100                 105                 110

Asp His Gln Tyr Tyr Val Tyr Phe Asp Ser Val Arg Asp Cys Tyr Glu
        115                 120                 125

Ala Phe Val Ile Tyr Ser Phe Leu Ser Leu Cys Phe Gln Tyr Leu Gly
    130                 135                 140

Gly Glu Gly Ala Ile Met Ala Glu Ile Arg Gly Lys Pro Ile Lys Ser
145                 150                 155                 160

Ser Cys Leu Tyr Gly Thr Cys Cys Leu Arg Gly Met Thr Tyr Ser Ile
                165                 170                 175

Gly Phe Leu Arg Phe Cys Lys Gln Ala Thr Leu Gln Phe Cys Leu Val
            180                 185                 190

Lys Pro Val Met Ala Val Thr Thr Ile Ile Leu Gln Ala Phe Gly Lys
        195                 200                 205

Tyr His Asp Gly Asp Phe Asn Val Arg Ser Gly Tyr Leu Tyr Val Thr
    210                 215                 220

Leu Ile Tyr Asn Ala Ser Val Ser Leu Ala Leu Tyr Ala Leu Phe Leu
225                 230                 235                 240

Phe Tyr Phe Thr Thr Arg Glu Leu Leu Arg Pro Phe Gln Pro Val Leu
                245                 250                 255

Lys Phe Leu Thr Ile Lys Ala Val Ile Phe Leu Ser Phe Trp Gln Gly
            260                 265                 270

Leu Leu Leu Ala Ile Leu Glu Arg Cys Gly Val Ile Pro Glu Val Glu
        275                 280                 285
```

```
Thr Ser Gly Gly Asn Lys Leu Gly Ala Gly Thr Leu Ala Ala Gly Tyr
    290                 295                 300

Gln Asn Phe Ile Ile Cys Val Glu Met Leu Phe Ala Ser Val Ala Leu
305                 310                 315                 320

Arg Tyr Ala Phe Pro Cys Gln Val Tyr Ala Glu Lys Lys Glu Asn Ser
                325                 330                 335

Pro Ala Pro Pro Ala Pro Met Gln Ser Ile Ser Ser Gly Ile Arg Glu
            340                 345                 350

Thr Val Ser Pro Gln Asp Ile Val Gln Asp Ala Ile His Asn Phe Ser
        355                 360                 365

Pro Ala Tyr Gln His Tyr Thr Gln Gln Ala Thr His Glu Ala Pro Arg
    370                 375                 380

Pro Gly Thr His Pro Ser Gly Gly Ser Gly Gly Ser Arg Lys Ser Arg
385                 390                 395                 400

Ser Leu Glu Lys Arg Met Leu Ile Pro Ser Glu Asp Leu
                405                 410
```

<210> 7

<211> 1317

<212> DNA

<213> Homo sapiens

<400> 7

```
atgccttgca cttgtacctg gaggaactgg agacagtgga ttcgaccttt agtagcggtc     60

atctacctgg tgtcaatagt ggttgcggtt cccctatgcg tgtgggaatt acagaaactg    120

gaggttggaa tacacaccaa ggcttggttt attgctggaa tctttttgct gttgactatt    180

cctatatcac tgtgggtgat attgcaacac ttagtgcatt atacacaacc tgaactacaa    240

aaaccaataa taaggattct ttggatggta cctatttaca gtttagatag ttggatagct    300

ttgaaatatc ccggaattgc aatatatgtg gatacctgca gagaatgcta tgaagcttat    360

gtaatttaca actttatggg attccttacc aattatctaa ctaaccggta tccaaatctg    420

gtattaatcc ttgaagccaa agatcaacag aaacatttcc ctcctttatg ttgctgtcca    480

ccatgggcta tgggagaagt attgctgttt aggtgcaaac taggtgtatt acagtacaca    540
```

gttgtcagac ctttcaccac catcgttgct ttaatctgtg agctgcttgg tatatatgac 600

gaagggaact ttagcttttc aaatgcttgg acttatttgg ttataataaa caacatgtca 660

cagttgtttg ccatgtattg tctcctgctc ttttataaag tactaaaaga agaactgagc 720

ccaatccaac ctgttggcaa atttctttgt gtaaggctgg tggtttttgt ttctttttgg 780

caagcagtag ttattgcttt gttggtaaaa gttggcgtta tttctgaaaa gcatacgtgg 840

gaatggcaaa ctgtagaagc tgtggccacc ggactccagg attttattat ctgtattgag 900

atgttcctcg ctgccattgc tcatcattac acattctcat ataaaccata tgtccaagaa 960

gcagaagagg gctcatgctt tgattccttt cttgccatgt gggatgtctc agatattaga 1020

gatgatattt ctgaacaagt aaggcatgtt ggacggacag tcgggggaca tcccaggaaa 1080

aaattgtttc ccgaggatca agatcaaaat gaacatacaa gtttattatc atcatcatca 1140

caagatgcaa tttccattgc ttcttctatg ccaccttcac ccatgggtca ctaccaaggg 1200

tttggacaca ctgtgactcc ccagactaca cctaccacag ctaagatatc tgatgaaatc 1260

cttagtgata ctataggaga gaaaaaagaa ccttcagata aatccgtgga ttcctga 1317

<210> 8

<211> 438

<212> PRT

<213> Homo sapiens

<400> 8

Met Pro Cys Thr Cys Thr Trp Arg Asn Trp Arg Gln Trp Ile Arg Pro
1 5 10 15

Leu Val Ala Val Ile Tyr Leu Val Ser Ile Val Val Ala Val Pro Leu
20 25 30

Cys Val Trp Glu Leu Gln Lys Leu Glu Val Gly Ile His Thr Lys Ala
35 40 45

Trp Phe Ile Ala Gly Ile Phe Leu Leu Leu Thr Ile Pro Ile Ser Leu
50 55 60

Trp Val Ile Leu Gln His Leu Val His Tyr Thr Gln Pro Glu Leu Gln
65 70 75 80

Lys Pro Ile Ile Arg Ile Leu Trp Met Val Pro Ile Tyr Ser Leu Asp

85 90 95

Ser Trp Ile Ala Leu Lys Tyr Pro Gly Ile Ala Ile Tyr Val Asp Thr
100 105 110

Cys Arg Glu Cys Tyr Glu Ala Tyr Val Ile Tyr Asn Phe Met Gly Phe
115 120 125

Leu Thr Asn Tyr Leu Thr Asn Arg Tyr Pro Asn Leu Val Leu Ile Leu
130 135 140

Glu Ala Lys Asp Gln Gln Lys His Phe Pro Pro Leu Cys Cys Cys Pro
145 150 155 160

Pro Trp Ala Met Gly Glu Val Leu Leu Phe Arg Cys Lys Leu Gly Val
165 170 175

Leu Gln Tyr Thr Val Val Arg Pro Phe Thr Thr Ile Val Ala Leu Ile
180 185 190

Cys Glu Leu Leu Gly Ile Tyr Asp Glu Gly Asn Phe Ser Phe Ser Asn
195 200 205

Ala Trp Thr Tyr Leu Val Ile Ile Asn Asn Met Ser Gln Leu Phe Ala
210 215 220

Met Tyr Cys Leu Leu Leu Phe Tyr Lys Val Leu Lys Glu Glu Leu Ser
225 230 235 240

Pro Ile Gln Pro Val Gly Lys Phe Leu Cys Val Arg Leu Val Val Phe
245 250 255

Val Ser Phe Trp Gln Ala Val Val Ile Ala Leu Leu Val Lys Val Gly
260 265 270

Val Ile Ser Glu Lys His Thr Trp Glu Trp Gln Thr Val Glu Ala Val
275 280 285

Ala Thr Gly Leu Gln Asp Phe Ile Ile Cys Ile Glu Met Phe Leu Ala
290 295 300

Ala Ile Ala His His Tyr Thr Phe Ser Tyr Lys Pro Tyr Val Gln Glu
305 310 315 320

Ala Glu Glu Gly Ser Cys Phe Asp Ser Phe Leu Ala Met Trp Asp Val
325 330 335

Ser Asp Ile Arg Asp Asp Ile Ser Glu Gln Val Arg His Val Gly Arg
340 345 350

Thr Val Gly Gly His Pro Arg Lys Lys Leu Phe Pro Glu Asp Gln Asp
355 360 365

Gln Asn Glu His Thr Ser Leu Leu Ser Ser Ser Ser Gln Asp Ala Ile
370 375 380

Ser Ile Ala Ser Ser Met Pro Pro Ser Pro Met Gly His Tyr Gln Gly
385 390 395 400

Phe Gly His Thr Val Thr Pro Gln Thr Thr Pro Thr Thr Ala Lys Ile
405 410 415

Ser Asp Glu Ile Leu Ser Asp Thr Ile Gly Glu Lys Lys Glu Pro Ser
420 425 430

Asp Lys Ser Val Asp Ser
435

<210> 9

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> hOST 1 PCR primer 1

<400> 9
atggagcagc ctgtgttcct gatg 24

<210> 10

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> hOST 1 PCR primer 2

<400> 10
ctagaattca tcatcagagc tgagc 25

<210> 11

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> hOST 2 PCR primer 1

<400> 11
atgagtaatg tctcagggat cctgg 25

<210> 12

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> hOST 2 PCR primer 2

<400> 12
ctacaggtcc tccgagggga tcagc 25

<210> 13

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> hOST 3 PCR primer 1

<400> 13
atgccttgca cttgtacctg gagg 24

<210> 14

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> hOST 3 PCR primer 2

<400> 14
tcaggaatcc acggatttat ctgaag 26

<210> 15

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> hOST 4 PCR primer 1

<400> 15
atggagccgg gcaggaccca gataa 25

<210> 16

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> hOST 4 PCR primer 2

<400> 16
ttaggctttg aggttcaagt ccagg 25

<210> 17

<211> 1228

<212> DNA

<213> Raja erinacea

<400> 17
cggacactca cagcgctggt cttaaccgcg cctgggatcc cgacactgcc aagaagatgg 60

atgtagctca ccctgaggaa gtgaccaggt tttctccaga tatcttgatg gaaaagttca 120

acgtttctga ggcgtgcttc ctgccccctc cgatatccat ccaactcata ctgcagctga 180

cgtggttaga cattggtgtc tttgccgcat tgaccgcgat gactgtgctc accatcgcca 240

tttacctgga gatcgtctgc tacctgatgg acaaggtgaa gtgtcccatc aagagaaaga 300

ctttgatgtg gaacagtgca gctccaaccg tcatcgccat cacttcctgc cttggtctct 360

gggtcccacg agccatcatg ttcgtggaca tggcggctgc catgtacttt ggtgttggct 420

tctacctgat gctgctgatc atcgtacagg ggtacggtgg agaggaggcc atgctccaac 480

acctggccac acacaccatc cgtatcagca ccgggccctg ctgctgctgc tgcccctgtc 540

taccccacat acacctcaca cggcagaaat acaagatctt tgtgctggga gctttccaag 600

tggctttcct ccggcctgcc ctcttcttgc tgggcgtggt cttgtggaca aacggcctct 660

atgacccaga tgattggtcc tccactagca tcttcctctg gctgaacctg ttcctgggcg 720

tttccaccat cctggggctg tggccggtca acgtcctctt ccgacactcc aaggtgctca 780

tggccgacca gaagctgacc tgcaagtttg ctctgttcca ggctatcctg atcctgtcct 840

cgctacagaa ttccatcatt ggaacgctgg cgggagcggg gcacattggc tgtgctcctc 900

cctattctgc aaggaccaga ggacagcaaa tgaacaacca gctgttgatt atcgagatgt 960

tcttcgttgg tatcctgacg cggatttcct acaggaagag ggatgaccga ccgggacacc 1020

gacatgtcgg tgaggtccag cagattgtca gagaatgtga tcaaccagcc atcgccgacc 1080

aacaggctga tcactccagc atctcccaca tataaacaca gatcagcaac ataactactt 1140

ggattaaaat gtgagatttt gcatggagaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1200

aaaaaaaaaa aaaaaaaaaa aaaaaaaa 1228

<210> 18

<211> 352

<212> PRT

<213> Raja erinacea

<400> 18

Met Asp Val Ala His Pro Glu Glu Val Thr Arg Phe Ser Pro Asp Ile

1 5 10 15

Leu Met Glu Lys Phe Asn Val Ser Glu Ala Cys Phe Leu Pro Pro Pro
20 25 30

Ile Ser Ile Gln Leu Ile Leu Gln Leu Thr Trp Leu Asp Ile Gly Val
35 40 45

Phe Ala Ala Leu Thr Ala Met Thr Val Leu Thr Ile Ala Ile Tyr Leu
50 55 60

Glu Ile Val Cys Tyr Leu Met Asp Lys Val Lys Cys Pro Ile Lys Arg
65 70 75 80

Lys Thr Leu Met Trp Asn Ser Ala Ala Pro Thr Val Ile Ala Ile Thr
85 90 95

Ser Cys Leu Gly Leu Trp Val Pro Arg Ala Ile Met Phe Val Asp Met
100 105 110

Ala Ala Ala Met Tyr Phe Gly Val Gly Phe Tyr Leu Met Leu Leu Ile
115 120 125

Ile Val Gln Gly Tyr Gly Gly Glu Glu Ala Met Leu Gln His Leu Ala
130 135 140

Thr His Thr Ile Arg Ile Ser Thr Gly Pro Cys Cys Cys Cys Cys Pro
145 150 155 160

Cys Leu Pro His Ile His Leu Thr Arg Gln Lys Tyr Lys Ile Phe Val
165 170 175

Leu Gly Ala Phe Gln Val Ala Phe Leu Arg Pro Ala Leu Phe Leu Leu
180 185 190

Gly Val Val Leu Trp Thr Asn Gly Leu Tyr Asp Pro Asp Asp Trp Ser
195 200 205

Ser Thr Ser Ile Phe Leu Trp Leu Asn Leu Phe Leu Gly Val Ser Thr
210 215 220

Ile Leu Gly Leu Trp Pro Val Asn Val Leu Phe Arg His Ser Lys Val
225 230 235 240

Leu Met Ala Asp Gln Lys Leu Thr Cys Lys Phe Ala Leu Phe Gln Ala

245 250 255

Ile Leu Ile Leu Ser Ser Leu Gln Asn Ser Ile Ile Gly Thr Leu Ala
260 265 270

Gly Ala Gly His Ile Gly Cys Ala Pro Pro Tyr Ser Ala Arg Thr Arg
275 280 285

Gly Gln Gln Met Asn Asn Gln Leu Leu Ile Ile Glu Met Phe Phe Val
290 295 300

Gly Ile Leu Thr Arg Ile Ser Tyr Arg Lys Arg Asp Asp Arg Pro Gly
305 310 315 320

His Arg His Val Gly Glu Val Gln Gln Ile Val Arg Glu Cys Asp Gln
325 330 335

Pro Ala Ile Ala Asp Gln Gln Ala Asp His Ser Ser Ile Ser His Ile
340 345 350

<210> 19

<211> 454

<212> DNA

<213> Homo sapiens

<400> 19
ctcgttgcac acgctaccag gagcaggggc atggagcaca gtgaggggggc tcccggagac 60

ccagccggta ctgtggtacc ccaggagctg ctggaagaga tgctttggtt ttttcgtgtg 120

gaagatgcat ctccctggaa tcattccatc cttgccctgg cagctgtggt ggtcattata 180

agcatggtcc tcctgggaag aagcatccag gcaagcagaa aagaaacgat gcagccacca 240

gaaaaagaaa ctccagaagt cctgcatttg gatgaggcca aggatcacaa cagcctaaac 300

aacctaagag aaactttgct ctcagaaaag ccaaacttgg cccaggtgga acttgagtta 360

aaagagagag atgtgctgtc agttttcctt ccggatgtac cagaaactga gagctagtga 420

gggttcagag aagccccatc ctaagccaga caca 454

<210> 20
<211> 128

<212> PRT

<213> Homo sapiens

<400> 20

```
Met Glu His Ser Glu Gly Ala Pro Gly Asp Pro Ala Gly Thr Val Val
1               5                  10                  15

Pro Gln Glu Leu Leu Glu Glu Met Leu Trp Phe Phe Arg Val Glu Asp
            20                  25                  30

Ala Ser Pro Trp Asn His Ser Ile Leu Ala Leu Ala Ala Val Val Val
        35                  40                  45

Ile Ile Ser Met Val Leu Leu Gly Arg Ser Ile Gln Ala Ser Arg Lys
    50                  55                  60

Glu Thr Met Gln Pro Pro Glu Lys Glu Thr Pro Glu Val Leu His Leu
65                  70                  75                  80

Asp Glu Ala Lys Asp His Asn Ser Leu Asn Asn Leu Arg Glu Thr Leu
                85                  90                  95

Leu Ser Glu Lys Pro Asn Leu Ala Gln Val Glu Leu Glu Leu Lys Glu
            100                 105                 110

Arg Asp Val Leu Ser Val Phe Leu Pro Asp Val Pro Glu Thr Glu Ser
        115                 120                 125
```

INTERNATIONAL SEARCH REPORT

PCT/US2005/002333

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/705 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data, PAJ, CHEM ABS Data, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DAWSON P.A. ET AL.: "The heteromeric organic solute transporter alpha-beta, Ostalpha-Ostbeta, is an ileal basolateral bile acid transporter." THE JOURNAL OF BIOLOGICAL CHEMISTRY. 25 FEB 2005, vol. 280, no. 8, 25 February 2005 (2005-02-25), pages 6960-6968, XP002328899 ISSN: 0021-9258 the whole document	1-12

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 May 2005	09/06/2005

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Bladier, C

INTERNATIONAL SEARCH REPORT

PCT/US2005/002333

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SEWARD D.J. ET AL.: "Functional complementation between a novel mammalian polygenic transport complex and an evolutionarily ancient organic solute transporter, OSTalpha-OSTbeta." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 278, no. 30, 25 July 2003 (2003-07-25), pages 27473-27482, XP002328898 ISSN: 0021-9258 the whole document & DATABASE EMBL Human organic solute transporter alpha, 340 aa 1 June 2003 (2003-06-01), retrieved from EBI Database accession no. Q86UW1 the whole document & DATABASE EMBL Organic solute transporter beta, 128 aa 1 June 2003 (2003-06-01), retrieved from EBI Database accession no. Q86UW2 the whole document	1-12
A	WANG W. ET AL.: "Expression cloning of two genes that together mediate organic solute and steroid transport in the liver of a marine vertebrate" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, vol. 98, no. 16, 31 July 2001 (2001-07-31), pages 9431-9436, XP002328897 ISSN: 0027-8424 cited in the application figures 2,3 page 9433, left-hand column, paragraph 3 - right-hand column, paragraph 1 & DATABASE EMBL 1228 bp 1 August 2001 (2001-08-01), "Raja erinacea organic solute transporter alpha mRNA" retrieved from EBI Database accession no. AY027664 the whole document & DATABASE EMBL 352 aa 1 August 2001 (2001-08-01), "Raja erinacea organic solute transporter alpha" retrieved from EBI Database accession no. AAK14805 the whole document	

-/--

INTERNATIONAL SEARCH REPORT

PCT/US2005/002333

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WANG W. ET AL.: "Expression cloning of a novel polygenic organic solute transporter (Ostalpha, Ostbeta) from the liver of a marine vertebrate" HEPATOLOGY, vol. 32, no. 4 Pt. 2, October 2000 (2000-10), page 316A, XP009048064 & 51ST ANNUAL MEETING AND POSTGRADUATE COURSES OF THE AMERICAN ASSOCIATION FOR THE STUDY OF LIVER DISE; DALLAS, TEXAS, USA; OCTOBER 27-31, 2000 ISSN: 0270-9139 the whole document	
A	WO 01/20331 A (XENOPORT, INC; DOWER, WILLIAM, J; GALLOP, MARK; BARRETT, RONALD, W; CU) 22 March 2001 (2001-03-22) cited in the application the whole document	

INTERNATIONAL SEARCH REPORT

PCT/US2005/002333

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 13-20
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.2

Claims Nos.: 13-20

- claim 13 relates to a conjugate comprising an agent linked to a conjugate moiety, wherein the conjugate is defined by reference to a desirable property, namely it has a Vmax of at least 1% of taurocholate for the transporter of SEQ ID N°2 co-expressed with the co-transporter of SEQ ID N°20, wherein the agent has a pharmaceutical activity without the conjugate moiety and the conjugate has a greater Vmax for the transporter than the agent without the conjugate moiety;

- claims 14-17 relate to a method of manufacturing or to the use of a conjugate comprising an agent linked to a conjugate moiety, wherein the conjugate is also defined by reference to a desirable property, namely its ability to be transported by the transporter of SEQ ID N°2 co-expressed with the co-transporter of SEQ ID N°20 with a higher Vmax than the agent alone;

- claims 18-20 relate to the use of a conjugate comprising an agent covalently linked to a conjugate moiety through linkage cleavable in vivo, wherein the conjugate is also defined by reference to a desirable property, namely its has a Vmax of at least 1% of taurocholate for the transporter of SEQ ID N°2 co-expressed with the co-transporter of SEQ ID N°20, wherein oral bioability of the agent due to the uptake of the conjugate is improved relative to the oral bioavaibility of the agent administered without the conjugate moiety;

The claims cover all conjugates having these properties, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for none. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the conjugate by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, no search has not been carried out for claims 13-20.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

out during examination before the EPO (see EPO Guideline C-VI, 8.5), should the problems which led to the Article 17(2) declaration be overcome.

INTERNATIONAL SEARCH REPORT

PCT/US2005/002333

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0120331 A	22-03-2001	AU 7703400 A	17-04-2001
		EP 1212619 A1	12-06-2002
		WO 0120331 A1	22-03-2001